THE ESTIMATED IMPACT OF THE TENNESSEE VOLUNTARY PREKINDERGARTEN PROGRAM ON CHILDREN'S LITERACY, LANGUAGE, AND MATHEMATICS SKILLS:

RESULTS FROM A REGRESSION-DISCONTINUITY DESIGN

RESEARCH REPORT

GEORGINE M. PION

MARK W. LIPSEY

PEABODY RESEARCH INSTITUTE

VANDERBILT UNIVERSITY

JUNE 2020

Recommended Citation

Pion, G. M., & Lipsey, M. W. (2020). *The Estimated Impact of the Tennessee Voluntary Prekindergarten Program on Children's Literacy, Language, and Mathematics Skills: Results from a Regression-Discontinuity Design*. Nashville, TN: Peabody Research Institute, Vanderbilt University.

Funding Source

The research reported here was supported by Grant #R305E090009 from the Institute of Education Sciences, U.S. Department of Education, for the study titled "Evaluation of the Effectiveness of Tennessee's Voluntary Pre-Kindergarten Program."

Disclaimer

The opinions and conclusions expressed in this report are those of the authors and do not represent the opinions and positions of the Institute of Education Sciences of the U. S. Department of Education.

Acknowledgements

Thanks for assistance from the Tennessee Department of Education and the individuals below without whom this study would not have been possible:

Connie Casha, Director, Office of Early Learning Programs, Middle Tennessee State University, and former Director of the Office of Early Learning, Division of Curriculum and Instruction.

Bobbi Lussier, Executive Director, Office of Professional Laboratory Experiences, Middle Tennessee State University and Former Assistant Commissioner of Special Populations.

Robert Taylor, Consultant and former Superintendent of Bradley County Schools.

Thanks to the former Tennessee's Consortium on Research, Evaluation, and Development (now the Tennessee Education Research Alliance) for assistance in obtaining and interpreting data from the Tennessee education data system.

Special thanks and acknowledgement go to the numerous members of the research team who have been invaluable to this project over many years: Dale Farran, Co-Principal Investigator; Kerry Hofer who played a key role in the design and implementation of the RDD; Carol Bilbrey, Janie Hughart, and Ilknur Sekmen who managed the data collection; the many data collectors who worked across the state; and Nianbo Dong who served as a data analyst for three years of the project.

Table of Contents

List of Tables and Figures

List of Appendix Tables

THE ESTIMATED IMPACT OF THE TENNESSEE VOLUNTARY PREKINDERGARTEN PROGRAM ON CHILDREN'S SCHOOL READINESS: RESULTS FROM A REGRESSION-DISCONTINUITY DESIGN

Introduction

This report describes the results of one of two major component of an evaluation of the Tennessee Voluntary Prekindergarten Program (TN-VPK), those being a randomized controlled trial (RCT) and a regression-discontinuity (RD) design. The primary goal of the RCT was to determine both short- and longer-term impacts of the statewide TN-VPK program on participating children. Of interest were not only TN-VPK effects on a range of outcome variables at the end of the pre-k year, but also the effects on academic achievement and related outcomes (e.g., grade retention, special education placements, and school disciplinary actions) during the early elementary years. The results of this part of the study through third grade are summarized in Lipsey, Farran, and Durkin (2018) as well as in several technical reports (e.g., Lipsey et al., 2013a, 2013b).

The RD component of this larger study was more limited in scope with a focus on assessing TN-VPK's impact on selected outcomes measured at the beginning of kindergarten. This design was not capable of assessing follow-up impact in later grades but was implemented with a probability sample that allowed generalization to TN-VPK classrooms statewide. Moreover, classroom observations and related data were collected for each sampled classroom to describe instructional features potentially related to variation in the impact of TN-VPK on the literacy, language, and mathematics skills of the children attending the respective pre-k classrooms. The results of this latter effort are the subject of another report; the present report focuses on the statewide impact estimates for the effects of TN-VPK on participating children as they began kindergarten.

A number of research studies have used the age-cutoff RD design that was utilized in the present study to estimate the impact of publicly funded pre-k programs such as TN-VPK (see Appendix P-2 for a summary). The popularity of this quasi-experimental design is partly a function of its intuitive appeal. It involves two groups of children and a pre-k program with an age cutoff for eligibility to enroll in the program during a given school year. In this situation, there are children near the age cutoff who are old enough to qualify and enroll and others who fall short, perhaps by only days, and are required to wait until the next school year to enroll. A year later, the first group (the treatment cohort) has completed pre-k and is just entering kindergarten. The second group (a waitlist control cohort) is just entering pre-k and thus has not yet experienced its effects. These two cohorts would generally be expected to differ only by age, and, at the age cutoff, the age difference is negligibly small. Comparing outcome measures a year after these children are sorted into these two school-year cohorts by the age cutoff thus provides an estimate of the effect of pre-k for children presumed otherwise to be substantially

similar. The logic of this comparison can be extended to children whose age is further away in both directions from the age cutoff by using age as a statistical control in the analysis on the assumption that only age has differentiated the two cohorts. To further match the children in these two cohorts, statistical controls with other variables can also be used, e.g., demographic characteristics. A fuller description of the pre-k age-cutoff RD design and the nature of the cohort differences that may occur and distort the results can be found in Lipsey, Weiland, Yoshikawa, Wilson, and Hofer (2015).

The first pre-k program evaluation to use the age-cutoff RD design investigated the effects of the program in Tulsa, Oklahoma (Gormley, Gayer, Phillips, & Dawson, 2005). Since the publication of that study, quite a number of age-cutoff RD studies of public pre-k programs have been conducted. Bibliographic citations and a summary of the findings of these studies are provided in Appendices P-1 and P-2. Included are evaluations of programs in Arkansas, California, Georgia, New Mexico, Michigan, New Jersey, North Carolina, Oklahoma, South Carolina, Virginia, and West Virginia; a county-wide program in Michigan; and city-wide programs in Boston and San Francisco, along with Tulsa. These studies have almost universally found positive effects on all or most of the outcomes assessed, especially on the cognitive skills most often examined.

The Tennessee Prekindergarten Program (TN-VPK)

Launched in 2005, the TN-VPK program is a full-day program that operates on the same calendar as the public school system in Tennessee. By statute, it gives priority for enrollment to 4-year-old children eligible for the federal free or reduced price lunch programs; second priority is given to other at-risk 4-year-old children, including English language learners, those in state custody or at risk due to circumstances of neglect or abuse, and other students who meet local at-risk criteria.

The program requires a licensed teacher and an aide in every classroom, a maximum of 20 children per class, and a curriculum chosen from a state-approved list. Contracts from the state for VPK classrooms are written with individual school districts. Some districts choose to partner with community agencies and/or Head Start, but the majority of the classes are housed in public schools. In 2008-09 when the evaluation reported here first began, nearly every school district in Tennessee offered at least one full-day VPK classroom with a total enrollment statewide of about 18,000 children.

Methods

In the age-cutoff version of the RD design, a pre-specified age cutoff for pre-k eligibility determines which children are eligible to enroll in the program for the coming school year. In the present study, children had to be four years old by September 30th of a given year and eligible for kindergarten the following year to enroll in TN-VPK. Children a year younger or less who were not four years old by that cutoff date had to wait until the next fall before they could enroll.

The outcome variables on which VPK participating and nonparticipating children were compared included six subtests of the Woodcock-Johnson III Achievement Battery (Woodcock, McGrew, & Mather, 2001). Children in the treatment (TN-VPK) cohort who had participated in TN-VPK the year before were assessed shortly after they entered kindergarten the following year. Their counterparts in the control (No Pre-k) cohort either had (a) applied to TN-VPK at the same time as the children in the treatment cohort but were not admitted because they were too young), or (b) waited to apply until the next year when they were age-eligible. This cohort of children (control cohort) also was assessed at the beginning of the next year, which for them was just shortly after they entered TN-VPK and thus before much exposure to the program.

Children in both the treatment and control cohorts came from classrooms chosen via a two-stage probability sampling strategy designed to represent TN-VPK classrooms throughout the state. The RD design was conducted over the course of four years from 2009-10 to 2012-13. Figure 1 presents a schematic of the overall evaluation design. Further details on the sample, procedures, outcome measures, and analysis are provided in the sections that follow.

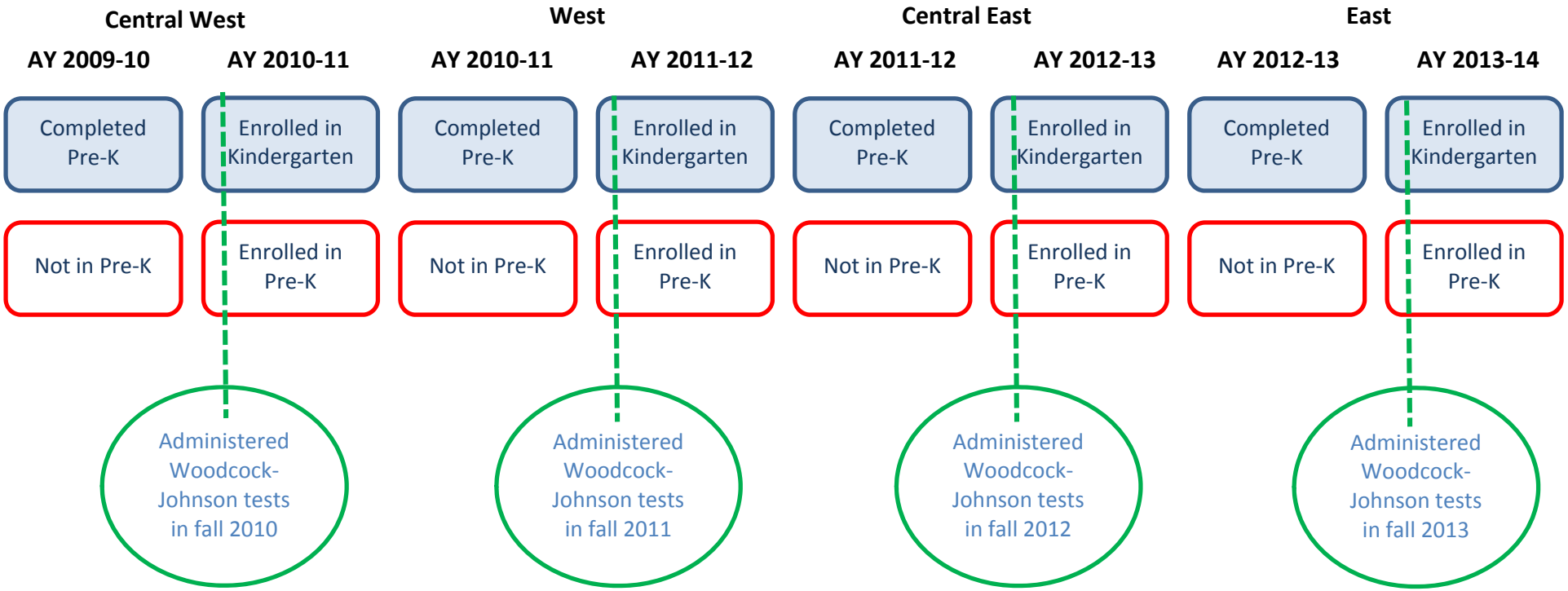
Sample

Implementation of the RD design began by selecting a stratified random sample of classrooms chosen to be representative of TN-VPK classrooms statewide. With assistance from the Tennessee Department of Education (TN-DOE) staff, four geographic regions (West, Central West, Central East, and East) were designated with the school districts and VPK programs within each identified. To be eligible for selection into the sample, each classroom that was part of the respective program had to have been in place for two successive school years and staffed by the same teacher both years. These criteria were imposed to increase the likelihood that the children in the two cohorts that would comprise the treatment and control groups were comparable with regard to their community or neighborhood locations and school preferences.

The 646 VPK programs operating across the regions at the time were described by their profiles on four key variables identified by TN-DOE staff as representing important distinctions between programs. These variables were (1) whether the program was in an urban or nonurban location; (2) whether it operated in the host school or in a partner community agency; (3) whether it was part of the original pilot VPK program or had been added when TN-VPK went to scale; and (4) whether it was located in a high priority host school (i.e., a school officially designated as among the lowest performing in the state). Based on these characteristics, the 942 TN-VPK classrooms found statewide within these 646 programs were assigned to a specific stratum, where a stratum consisted of a combination of these four characteristics, e.g., urban location, operated in a school, post-pilot program, and in a priority school.

Disproportionate random sampling of one classroom per selected program from these regions and strata was then conducted with sampling fractions created to sample larger proportions of classrooms from small strata than from large strata to ensure that smaller strata contributed sufficient numbers of classrooms to the total sample to allow them to be adequately represented in the analysis.

Figure 1
Structure of the Regression-Discontinuity Design



Note. AY refers to Academic Year. The blue outlined and shaded boxes represent the TN-VPK or treatment cohort; the red outlined and unshaded boxes refer to the No Pre-K control cohort.

Across the four regions, this sampling strategy resulted in the selection of 155 eligible classrooms located in 154 TN-VPK programs (one atypical program contributed two classrooms). Five classrooms from the original sample of 160 were dropped for various reasons. The population of children in one pre-k classroom changed from 4-year-olds to 3-year-olds over the course of the two study years. One classroom was discontinued after the end of the first study year, and another classroom moved to a nonparticipating school. Finally, there were two instances involving a teacher moving to a different school in the second study year—a school that served a much different population of children. The programs contributing classrooms were spread across 73 school districts and 62 counties. Sampling weights were assigned to each classroom so that the data from the sampled classrooms could be weighted back to the original proportions of the full statewide population of classrooms. These weights were calculated as the overall sampling fraction (155/942 =.1645) divided by the sampling fraction for the respective stratum (number of sampled classrooms in the stratum divided by the total number of classrooms in that stratum). Appendix A provides a more detailed description of the sampling procedure. As a practical matter, it was not possible to collect the child-level and classroom-level data required by the design for all the sampled classrooms within the same school year. The data collection therefore was spread over four school years with the sampled classrooms within one of the four regions participating in that process each year (as shown in Figure 1 above).

Student Participant Criteria

Once the sample of classrooms was chosen, the children enrolled in the selected classrooms were identified. The region chosen for the first wave of data collection was Central West, the region in which our research team was located. This meant that the TN-VPK (treatment) cohort in this region began pre-k in the fall of 2009. After they completed that pre-k year, they were then assessed within the first six weeks of kindergarten the next fall as long as they were still in the Tennessee public school system (see Figure 1). The corresponding entry years for treatment group pre-k participation in the other three regions were fall 2010 (West), fall 2011 (Central East), and fall 2012 (East).

The control group consisted of the cohort of children entering those same pre-k classrooms in the fall of the following year—2010 for Central West, 2011 for West, 2012 for Central East, and 2013 for East. These were the children who had missed the age cutoff for the prior year but then enrolled the next fall. They were assessed within the first six weeks after they began pre-k.

In order to make the two cohorts as comparable as possible, identical eligibility criteria were applied to select the children to be included in the analysis for both the treatment and control cohorts. These criteria were as follows:

- Children met the age-eligible requirement imposed by Tennessee. That is, a child was eligible to enroll in a TN-VPK classroom if he or she turned 4 years old by September 30th of the respective school year. Date of birth was determined from teacher records; if a child had a recorded date of birth that looked questionable (either much younger or older than the sample mean), we double-checked with the teacher that the date was correct.
- Children were enrolled in a participating TN-VPK classroom within the first weeks of their system-specific school start date for the appropriate pre-k year. Each child's enrollment date was determined from teacher reports.

- Children stayed in the same participating TN-VPK classroom in which they began the year until at least the last 6 weeks of their system-specific school end date.[1] Withdrawal dates were determined from teacher reports, information obtained during the assessment period, and/or classroom rosters obtained at the end of each pre-k year.
- Children were located in a Tennessee public school in the year following their pre-k year within the first six weeks of their system-specific school start date for the appropriate kindergarten year.[2] Notably, this meant following the control cohort of children for the year after their initial VPK enrollment to determine if they were then enrolled as kindergarteners in a Tennessee public school.
- Children did not withdraw from school in their kindergarten year within the first six weeks of the school year. This helped ensure that the treatment children were available to be assessed and the control children were comparable with regard to their kindergarten participation.

Application of these criteria resulted in some children being excluded from the analysis sample. For example, some children withdrew from pre-k early, enrolled too late, or moved to a classroom that was not one of the sampled classrooms.

Overall, approximately 16% of the children in the 155 sampled classrooms became ineligible for one or more of these reasons; this included 17% of treatment and 15% of control cohort children. Figure 2 diagrams the flow of children through the study.

After application of the above criteria and classroom deletions, the final sample included 5,189 children from successive cohorts moving through 155 classrooms. There were 2,622 children in the treatment cohort and 2,567 in the control cohort. On average, there were 17 treatment children and 17 control children per classroom (see Table 1).[3]

Table 1
Numbers of Children per Participating Classroom by Group

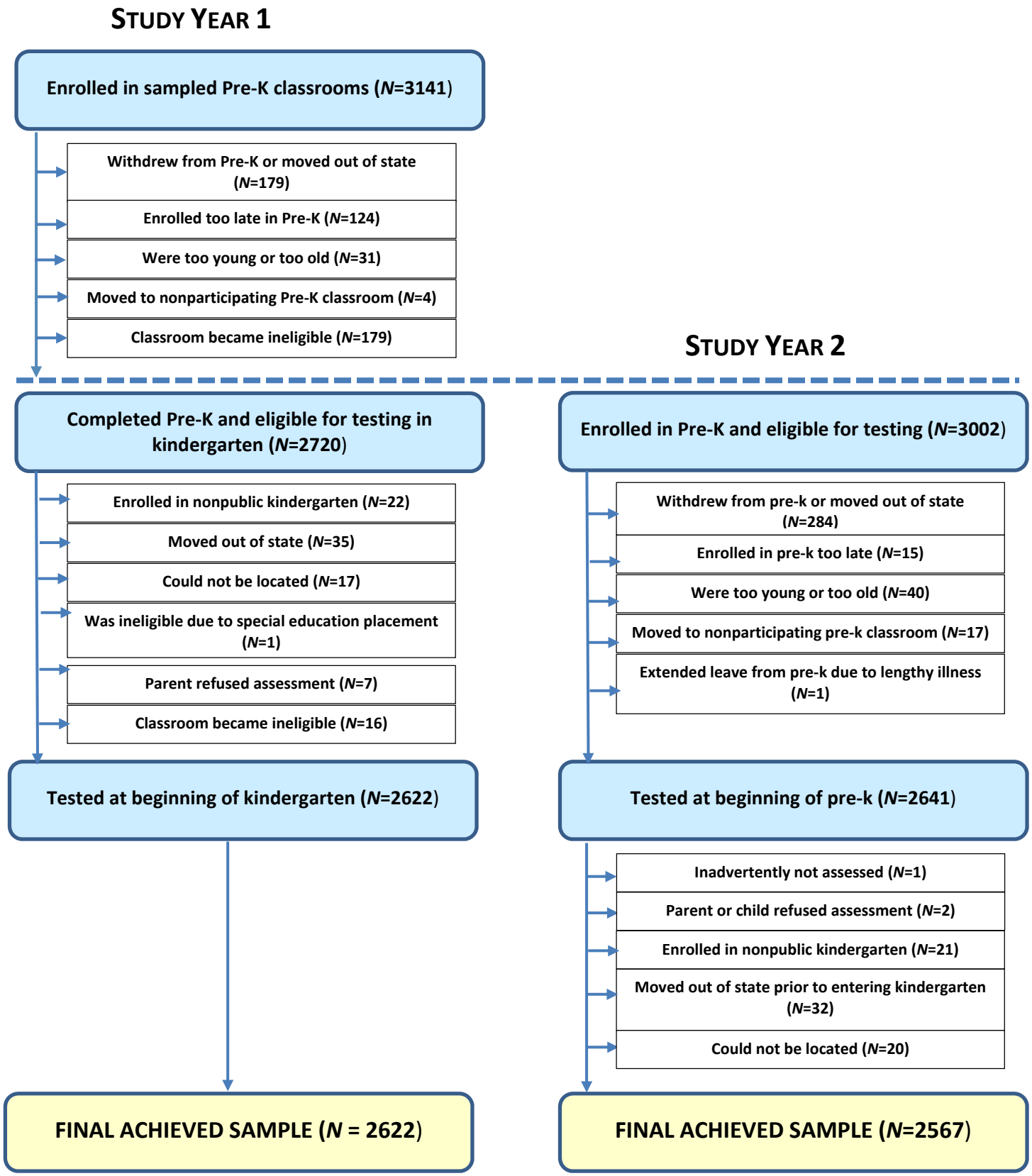
Group	Minimum	Maximum	Mean	SD	*N* of Classrooms
TN-VPK (Treatment)	10	20	16.9	2.17	155
Non-VPK (Control)	6	20	16.6	2.68	155

[1] In 5 instances, a student enrolled in one participating TN-VPK classroom but withdrew within the 1st 6 weeks of school and enrolled in another participating TN-VPK classroom; those students were matched with the latter classroom.

[2] We did not necessarily exclude children from study participation if they were not in kindergarten during the supposed kindergarten year. Although most students were promoted to kindergarten, occasionally a child was retained in pre-k or promoted early to 1st grade. Such students remained in our study sample.

[3] The large majority (83%) of the 155 classrooms had the same teacher for both study years. The exceptions involved the same classroom in the same school, but with a new teacher assigned for the second (control) cohort of children the second year: these classrooms were retained in the study inasmuch as the children in the control group were assessed at the beginning of the pre-k year prior to much exposure to that different teacher.

Figure 2
Consort Chart for the Achieved Sample of Children



Outcome Measures

Determining the effect of TN-VPK on children's academic skills at the time of entry into kindergarten was a major focus of the study. To accomplish this, children were administered a selection of measures from the Woodcock-Johnson III Achievement Battery (McGrew, Schrank, & Woodcock, 2007; Woodcock, McGrew, & Mather, 2001), a nationally normed measure often used in studies of the effects of early childhood education. Consistent with other age-cutoff RD studies of public pre-k programs (e.g., Gormley, Gayer, Phillips, & Dawson, 2005; Weiland & Yoshikawa, 2013), the following subscales were selected as outcome measures:

- Early Literacy
  - *Letter-Word Identification*: Ability to identify and pronounce alphabet letters and read words by sight.
  - *Spelling*: Ability to trace simple shapes and write orally presented letters and words correctly.
- Language
  - *Oral Comprehension*: Ability to listen to and provide a missing key word to an orally presented passage.
  - *Picture Vocabulary*: Extent of expressive vocabulary.
- Early Mathematics
  - *Applied Problems*: Ability to identify the necessary information and apply the appropriate strategy to solve numerical and spatial problems accompanied by pictures.
  - *Quantitative Concepts*: Ability to understand number identification, sequencing, shapes, symbols, and sequencing of numbers.

An overall index of achievement (WJ Composite) also was constructed by averaging a child's scores across these six subscales.

In our analyses, we used *W* scores, which have two useful psychometric properties. First, the *W* scale is an equal-interval scale; this means that an increase in *W* units on any of the tests can be interpreted in the same way. Second, the *W* scale was constructed using Rasch modeling, thus any *W* score represents both a child's ability level in a given area and the difficulty levels of the items. Because of these characteristics, the *W* scores are viewed by the test's authors as those most suitable for statistical calculations (Jaffe, 2009).

Data Collection Procedures

Children in the VPK treatment cohort were directly assessed within the first six weeks of their supposed kindergarten year (regardless of actual grade), and children in the control cohort were directly assessed within the first six weeks of their pre-k year. Trained assessors administered the WJ subtests to all students at their classroom sites. All subtests were administered in English, the language of instruction for all pre-k classrooms.

Basic demographic information for all children was obtained from records made available by the Tennessee Department of Education. These data included each child's age, gender, race/ethnicity, and native language as well as eligibility for free or reduced price lunch and whether the child had an

Individual Educational Plan (IEP) during the pre-k year. We also collected separate data on the child's age and race/ethnicity during the assessment process.

Final Analytic Sample

Of the 5,189 children who met the original criteria for our sample, school records identified 4,144 (79.9%) as eligible for the federal free or reduced price lunch programs (FRPL); this included 2,078 children in the treatment cohort and 2,066 children in the control cohort. The primary analytic sample was restricted to these children for two reasons. First, these economically disadvantaged children constituted the primary target population for TN-VPK. Second, the analytic sample of the RCT component of the broader TN-VPK research study only included FRPL-eligible children. Thus, imposing this criterion on the RD analytic sample facilitated comparison of the results from the two separate study components. This decision did result in the number of classrooms represented in the sample decreasing from 155 to 151.[4]

Missing Data

There were no missing data on the age, gender, and race/ethnicity of the children in the analytic sample. Data on a child's native language, eligibility for free or reduced lunch, and special education placement were obtained from the state's education data system. For any child, all data records on these characteristics for their pre-k, kindergarten, and first grade years were extracted. In those cases where information on a variable (e.g., native language) was missing for a child's pre-k year, it often was reported for later years. This information was then used for their pre-k standing. This strategy resulted in an almost complete data set for use in analysis. The only variable for which there were missing data was free or reduced lunch eligibility (recorded as free lunch or reduced lunch eligible or eligible for neither), and only one child had no information on FRPL status.

Data Analytic Strategy

In order for the RD design to produce valid estimates of treatment effects, at least three conditions must be satisfied (Jacob, Zhu, Somers, & Bloom, 2012). First, there must be a clear cut point on the assignment variable, i.e., distinct values on some variable that determines whether a child is in the treatment or control group. Second, there should be no differences between the treatment and control groups across the cut point on any characteristics other than treatment vs. control status that might be related to the outcomes. Finally, the functional relationship between the assignment variable and the outcome variable must be correctly specified. Each of these is examined below for the TN-VPK data.

A clear assignment variable. In this study the exogenous determinant of children's eligibility for TN-VPK was age. As previously noted, children in our sample must have turned 4 years of age by September 30th in order to attend TN-VPK. Even one day on either side of that cutoff was defined by the Tennessee Department of Education as sufficient to differentiate eligible from ineligible children at the beginning of any given school year.

Equivalence between the TN-VPK (treatment) and no pre-k (control) groups. Our application of the age-cutoff RD design required that children in both cohorts enroll in the same pre-k

[4] Three (7.6%) of the 39 classrooms in the Central East region and 1 (2.5%) of the 39 classrooms in the East region had no FRPL-eligible children enrolled during the study years.

classrooms in the same schools, as well as in a Tennessee public school kindergarten classroom the next year, in order to be represented in the sample. The importance of this design element is that it should increase the probability that the children in the two age cohorts for each school would be comparable on many characteristics related to the local context and the catchment areas of the respective VPK programs.

For the most part, these design requirements were adhered to when constructing the sampling frame and selecting the analytic sample, but there were a few exceptions. One of the 154 schools in the sample had two classrooms, both of which were included in the study. In addition, although all classrooms had the same teacher for children for both age cohorts at the time of sample selection, there were instances of unexpected personnel changes that occurred prior to the testing of the children. This involved 25 (16.6%) of the 151 classrooms; these classrooms were kept in the study.

As noted earlier, the demographic characteristics for children obtained from the Tennessee Department of Education included gender, race/ethnicity, primary language spoken at the child's home, and whether the child had an Individualized Education Plan (IEP) in the pre-k year. To compare the cohorts on these variables, each was used as the dependent variable in a multilevel model with children nested within classrooms and condition (treatment or control cohorts) as the only predictor.[5] These multilevel models were repeated for each variable with three different age bandwidths—children born within ±12 months of the September 30th eligibility cutoff date, those born within ±6 months of this date, and those born within ±3 months of this date. These bandwidths were chosen so that the results could be easily compared with previous studies (e.g., Gormley et al., 2005; Weiland & Yoshikawa, 2013).

Table 2 reports the results of these analyses for the three bandwidths. For the 12-month bandwidth, the two cohorts were similar with regard to gender, race, language spoken at home, and IEP placement. Compared to the treatment cohort, the control cohort did have a slightly larger proportion of Hispanic children (.054 vs. .044), but the difference was only marginally significant. The TN-VPK cohort was significantly older, consistent with the function of age as the key assignment variable for the treatment and control cohorts. The two cohorts also differed on the timing of the outcome assessments, with the time between the start of school and administration of the WJ subtests being a few days longer on average for the treatment cohort. This was due to the somewhat different assessment circumstances for the two cohorts. Children in the No Pre-k control cohort were all located in their respective pre-k classrooms. In contrast, the kindergarten classrooms in which the TN-VPK treatment cohort children were enrolled had to be identified and located. Although teachers provided this information to the best of their knowledge at the end of children's pre-k year, it was not always accurate and required extra time to update. In addition, some children were in kindergarten classrooms in different schools than those where their TN-VPK classrooms were located, which then increased the scheduling and travel time for assessors.

[5] The state administrative data included separate variables for race and ethnicity, and children could be members of multiple racial groups. In the analytic sample, 70.1% of children were White, 30.5% were Black, and 8.8% were Hispanic. The percentages of children who were Native American or Asian/Pacific Islander were quite small (0.4% and 0.8%, respectively). As such, three dummy variables were used in the baseline analyses to represent a child's race and/or ethnicity: whether the child was Black, Hispanic, or White respectively. Also, because of children's young ages, formal designation as English Language Learners (ELL) was either not done in some districts or was unreliable. Thus, we relied on data about the primary language spoken at home to characterize English language proficiency.

Table 2
Comparison of TN-VPK and No Pre-K Cohorts on Baseline Characteristics: 12-, 6-, and 3-Month Bandwidth Samples

Variable	Means[a]		Odds Ratio	TN-VPK vs. No Pre-K Difference (Pooled SD)	Effect Size[b]	*p*-value
	TN-VPK	No Pre-K				
12-month bandwidth sample	(*N*=2078)	(*N*=2066)				
Gender (Male)	.495	.494	1.01		.004	.915
White	.788	.809	.88		-.073	.220
Black	.216	.201	1.10		.052	.368
Hispanic	.044	.054	.80		-.127	.063
Primary language is not English	.032	.027	1.19		.095	.221
IEP	.090	.084	1.07		.039	.512
Age at testing (months)	65.5	53.4		12.08 (6.93)	1.743	.001
Days elapsed between start of school and testing	35.2	32.8		2.41 (5.72)	.421	.001
6-month bandwidth sample	(*N*=1099)	(*N*=1058)				
Gender (Male)	.493	.477	1.07		.037	.462
White	.782	.810	.84		-.096	.262
Black	.236	.196	1.27		.132	.110
Hispanic	.047	.056	.83		-.103	.326
Primary language is not English	.038	.043	.89		-.064	.560
IEP	.096	.078	1.26		.128	.144
Age at testing (months)	62.4	56.4		5.97 (3.42)	1.747	.001
Days elapsed between start of school and testing	35.2	33.1		2.12 (5.48)	.387	.001
3-month bandwidth sample	(*N*=519)	(*N*=531)				
Gender (Male)	.484	.496	.95		-.028	.695
White	.726	.783	.73		-.174	.114
Black	.292	.216	1.50		.224	.037
Hispanic	.037	.074	.49		-.394	.010
Primary language is not English	.032	.060	.51		-.372	.024
IEP	.091	.089	1.03		.016	.882
Age at testing (months)	60.9	57.8		3.10 (1.75)	1.771	.001
Days elapsed between start of school and testing	36.2	32.3		3.97 (5.32)	.746	.001

Note. Only children identified as eligible for free or reduced lunch were included in the analysis. Analyses were weighted to project the estimates to the statewide population of VPK classrooms.

[a] Reported are estimated means from the multilevel models with children nested within schools, membership in the treatment group as the only predictor with the respective baseline variable as the dependent variable.

[b] For binary variables (gender, white, black, Hispanic, primary language, and IEP), odds ratios were converted into effect sizes, using Chin's (2000) method. For continuous variables (age and days elapsed since start of school and testing), effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the pooled standard deviation of the two groups.

For the 6-month bandwidth, the results were similar to those for the full 12-month analytic sample; there were no significant differences between the TN-VPK and No Pre-k cohorts with the exception of age and the days between the start of the school year and administration of the WJ subtests. In contrast, the 3-

month bandwidth sample differed not only on these two variables, but also on race/ethnicity and primary language spoken at home. The proportion of Black children was higher in the treatment cohort whereas the proportions of Hispanic children and those from non-English speaking households were larger in the control cohort.

These demographic variables were included as covariates in all analyses of VPK effects on the various outcome measures in order to adjust for whatever differences occurred between the cohorts on them.

Accurate specification of the functional relationship between age and the outcome variables. To determine the appropriate functional form, the strategies outlined by Jacob et al. (2012) were used. First, graphical depictions of the relationship between each Woodcock Johnson measure and age (i.e., the number of days before and after the September 30th date for eligibility) were examined. This provided a visual assessment of whether a discontinuity at the cut point was evident and whether there was any indication of discontinuities at points other than the cut point (Lee & Lemieux, 2010). For each outcome variable, a smoothed plot of the mean of the outcome at the midpoint value of the age-based assignment variable was constructed for equal-sized intervals. The appropriate interval size—one that did not under- or over-smooth the relationships—was determined, using two formal statistical tests (Jacobs et al., 2012).

Second, a variety of functional forms (e.g., linear, linear interaction, and cubic) for the relationships between the assignment variable (age) and the various outcome variables were tested to determine which fit the data best. Following the advice of Lee and Lemieux (2010) and Jacob et al. (2012), two regressions were run for each outcome. The first involved regressing the outcome on treatment condition (TN-VPK treatment versus control), the assignment variable, and the appropriate term(s) for the functional form being tested. For the assignment variable, we created a new variable that measured the number of days from when the child turned 4 years old, centered on the cutoff of September 30th for the respective year. The second regression again involved these variables as well as indicator variables that represented the set of intervals used in the graphical displays described in the previous paragraph. The F-test approach outlined by Jacob et al. (2012) was used to compare the two models; when the results were not statistically significant, this was interpreted as indicating that no unexplained variability in the simple model existed that was captured by the more complex model.

Estimation of the TN-VPK effect. All analyses were conducted with multilevel regression models with children nested within schools. Variables included in each model were:

- *Age*. Age was centered at the cutoff for pre-k eligibility with children whose birthdays fell before September 30th receiving positive values according to how many days before that date. Those with birthdays after that date received corresponding negative values.
- *Condition*. A "1" was assigned to children who had completed TN-VPK and were enrolled in TN public schools during the kindergarten year, with a "0" being given to children who had just begun TN-VPK, the treatment and control cohorts respectively.
- *Region*. Because the sampling design had region as a stratification variable with data collected over successive years, region of the pre-k classroom served as a fixed effects blocking factor in the design. Three dummy variables, one each for the West, Central East, and East regions, were used to represent these four regions with Central West omitted as the reference value.

- *Covariates*. In order to adjust for baseline differences (see Table 2 above) and improve statistical power, both the baseline demographic characteristics and the timing-of-assessment variable were included as covariates. Demographic characteristics included: gender (male=1, female=0); whether the child was Black (Yes=1, No=0); whether the child was Hispanic (Yes=1, No=0); whether a language other than English was the primary language spoken at home (Yes=1, No=0); and whether the child had an IEP (Yes=1, No=0). The number of days between administration of the Woodcock Johnson subtests and the start of the school year also was included and centered at the grand mean.

These multilevel models were run for each of the bandwidths: ±12 months, ±6 months, and ±3 months around the cutoff date for TN-VPK eligibility. To check the robustness of the results of these analyses, pre-k effects were re-estimated after excluding the outermost 1%, 5%, and 10% of the data points with the highest and lowest age values. If the functional form was not misspecified, the impact estimates should not markedly change as a result of dropping these outermost data points (Jacob et al., 2012).

Estimation of effects for specific subgroups. A relevant question is whether the effects of TN-VPK were larger or smaller for certain subgroups of children. Of particular interest was whether there were differences in the effects on any of the various outcome measures for males versus females, Hispanics and Blacks versus Whites, and children from homes where the primary language was English versus their counterparts where Spanish or another foreign language was primarily spoken at home.[6] Multilevel models similar to those described for the full analytic sample were repeated for the Woodcock Johnson composite measure while including the respective subgroup membership variable and the term for its interaction with the treatment condition variable.

Results

Estimated Effects on Literacy, Language, and Math Skills

Participation in the TN-VPK program showed statistically significant effects on five of the six Woodcock-Johnson subtests plus the WJ composite of all six subtests.[7] As shown in Table 3, the effect sizes for the overall composite were large--.85, .83, and .91 for the 12-, 6-, and 3-month bandwidths, respectively. The greatest improvement was found for early literacy skills where the effect sizes for Letter-Word and Spelling were close to or slightly exceeded one standard deviation.

The effect sizes for emergent mathematics skills were more modest but still noteworthy. For the Applied Problems and Quantitative Concepts subtests, they were .71 and .60 for the 3-month bandwidth and ranged between .47 and .56 for the 6- and 12-month bandwidths.

[6] Twenty-seven different languages were represented among the nearly 8% of the analytic sample where the primary home language was not English. The most frequent language spoken was Spanish (83.6%), with the remaining languages each accounting for 2% or less.

[7] Results of the full models for the WJ composite measure and each Woodcock-Johnson subtest and each bandwidth are reported in Appendix Tables B-1 to B-7.

Table 3
Estimates of the Pre-K Effect for the WJ Composite and Subtests

Outcome and Bandwidth	Means[a]		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K Difference[b]	*p*	Effect size[c]
	TN-VPK	No Pre-K				
Composite of 6 subtests						
± 12 months	411.9	401.0	12.92	10.93	.001	.85
± 6 months	412.0	400.2	12.66	10.74	.001	.83
± 3 months	411.9	400.8	11.94	10.84	.001	.91
Letter Word						
± 12 months	346.7	324.7	21.15	22.02	.001	1.04
± 6 months	346.8	325.0	19.71	21.83	.001	1.11
± 3 months	345.4	325.6	19.17	19.85	.001	1.04
Spelling						
± 12 months	379.5	359.2	20.82	20.34	.001	.98
± 6 months	379.7	360.4	20.62	19.28	.001	.94
± 3 months	379.3	359.5	20.43	19.76	.001	.97
Oral Comprehension						
± 12 months	451.5	448.3	13.71	3.25	.001	.24
± 6 months	451.3	448.4	13.08	2.91	.009	.22
± 3 months	452.1	448.1	12.87	4.03	.012	.31
Picture Vocabulary						
± 12 months	464.8	461.1	10.72	3.69	.001	.34
± 6 months	464.9	461.2	10.89	3.69	.002	.34
± 3 months	464.5	461.8	10.41	2.69	.107	.26
Applied Problems						
± 12 months	408.0	399.7	17.74	8.37	.001	.47
± 6 months	409.4	399.6	18.12	9.80	.001	.54
± 3 months	409.8	398.4	16.08	11.36	.001	.71
Quantitative Concepts						
± 12 months	420.7	412.8	13.99	7.88	.001	.56
± 6 months	419.9	412.8	13.43	7.08	.001	.53
± 3 months	419.7	412.3	12.49	7.45	.001	.60

Note. The sample sizes for the TN-VPK and No Pre-K groups were: 2,078 and 2,066 for ±12 months; 1009 and 1,058 for ±6 months; and 519 and 531 for ±3 months. Only children eligible for free or reduced price lunch were included. All analyses were weighted to project the estimates to the statewide population of TN-VPK classrooms.

[a] Reported are estimated marginal means from the multilevel analysis model.

[b] Estimates were based on a multilevel model with children nested within pre-k classroom. In addition to condition and days from the eligibility cutoff (centered at zero), covariates included: (1) region; (2) whether the child was male; (3) whether the child was Black; (4) whether the child was Hispanic; (5) whether the child's native language was not English; (6) whether the child had a special education placement (IEP); and (7) the number of days between the WJ testing date and the start of school (centered at the grand mean). Previous analyses showed that the appropriate functional form for all models was linear. Appendix Tables B1 to B7 present the results for the full models for each outcome and bandwidth sample.

[c] Standardized mean difference effect sizes calculated by dividing the TN-VPK vs. No Pre-K difference by the standard deviation of the TN-VPK treatment group, recognizing that group as more analogous to the posttest outcomes typically used to compute effect sizes.

The smallest gains occurred in children's language skills. Effect sizes ranged from .24 to .31 for Oral Comprehension and from .26 to .34 for Picture Vocabulary. With regard to the latter, the estimate of the TN-VPK effect for children in the 3-month bandwidth was only marginally significant.

In order to assess the robustness of these results, each model was re-estimated on samples that dropped the outermost 1%, 5%, and 10% of the youngest and oldest children. The resulting TN-VPK vs. No Pre-K coefficients for the effect estimates were compared by standardizing their difference by dividing by the standard deviation of the TN-VPK cohort, corresponding to the difference in effect size estimates derived from the model with all the cases and the respective model with cases trimmed. As an overall summary, the absolute values of these effect size differences were averaged across the seven WJ outcome measures for each level of trimming and each bandwidth (see Appendix Table C-1).

For the 12-month sample vs. the most trimmed sample, the summary values showed that the average effect size difference was .019. These summary values were larger for the 6- and 3-month samples, .078 and .084, respectively, but were still relatively small (shown in the last rows of Appendix Table C-1). These results lend additional confidence that the models captured the correct functional form of the relationship between age and the WJ subtest scores.

Estimated Effects for Selected Subgroups

Although the overall effects of TN-VPK on children's achievement skills were positive and mostly moderate to large in magnitude, the effects may have varied across different subgroups of children. This question was investigated by comparing outcomes by gender, race/ethnicity (Hispanics versus non-Hispanics and Blacks versus other racial groups), and primary language (English versus another language).[8] Similar to the previous strategy used for examining the main effects of TN-VPK, multilevel models were estimated for the 12-, 6-, and 3-month bandwidths using age, treatment condition, and the same set of covariates, but with interaction terms for the TN-VPK vs. No Pre-K treatment condition and the respective subgroup of interest.[9]

Statistically significant differential TN-VPK effects were found for only two of the subgroup comparisons—Hispanic children compared to their non-Hispanic counterparts, and children from households where English was vs. was not the primary language. The results for the overall WJ composite measure for these two subgroup comparisons are summarized in Table 4.[10] For each subgroup and each bandwidth, the TN-VPK and No Pre-K covariate-adjusted group means are reported with the corresponding effect sizes and the statistical significance for the interaction terms.

[8] The numbers of children from other racial/ethnic groups (e.g., Asian, Pacific Islander, and Native American) were too small to allow separate comparisons. A similar situation existed with regard to primary language. English was the primary language for approximately 92% of children; with the exception of Spanish (which accounted for 6%, each of the other 26 languages represented accounted for .2% or less. It also should be noted that because school districts had varying policies about making ELL designations for children at this young age, we used primary language to represent this characteristic.

[9] The full results of these multilevel models are reported in Appendix Tables D1 to K7.

[10] Although no significant differences were found on the WJ composite, there was some suggestion that female children scored higher on Oral Comprehension and that Black children did better than non-Black children on the Picture Vocabulary subtests.

Table 4

TN-VPK Effect Estimates on the WJ Composite for Hispanic and Native Language Subgroups of Children

Outcome and Bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
12-month bandwidth								
Ethnicity					12.92			<.001
Hispanic	406.5	168	386.8	186		19.72	1.53	
Not Hispanic	412.4	1910	402.3	1880		10.15	.79	
Primary language								<.001
English	412.0	1914	401.9	1926		10.07	.78	
Another language	410.1	164	389.0	140		21.12	1.64	
Ethnicity and primary language								<.001
Hispanic and English	408.5	49	387.3	72		21.18	1.64	
Hispanic and another language	398.8	119	379.7	114		19.10	1.48	
Not Hispanic and English	412.8	1865	403.0	1854		9.81	.76	
Not Hispanic and another language	413.8	45	384.7	26		29.12	2.25	
6-month bandwidth								
Ethnicity					12.66			<.001
Hispanic	405.1	78	384.8	87		20.31	1.60	
Not Hispanic	412.6	931	402.7	971		9.97	.79	
Primary language								<.001
English	412.0	936	402.1	980		9.90	.78	
Another language	411.3	73	390.3	78		21.02	1.66	
Ethnicity and primary language								<.001
Hispanic and English	407.8	25	381.9	27		25.89	2.05	
Hispanic and another language	397.7	53	379.9	60		17.81	1.41	
Not Hispanic and English	412.9	911	403.4	953		9.53	.75	
Not Hispanic and another language	416.3	20	384.6	18		31.70	2.50	

Table 4 (*continued*)

TN-VPK Effect Estimates on the WJ Composite for Hispanic and Native Language Subgroups of Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
3-month bandwidth								
Ethnicity					11.94			.007
Hispanic	404.1	35	385.0	41		19.08	1.60	
Not Hispanic	412.5	484	402.2	490		10.26	.86	
Primary language								.003
English	412.0	487	402.0	492		10.01	.84	
Another language	409.1	32	389.0	39		20.13	1.69	
Ethnicity and primary language								.440
Hispanic and English	401.0	11	385.0	12		15.99	1.34	
Hispanic and another language	396.7	24	376.8	29		19.90	1.67	
Not Hispanic and English	413.0	476	403.1	480		9.92	.83	
Not Hispanic and another language	408.2	8	386.6	10		21.52	1.80	

Note. Only children eligible for free or reduced price lunch were included. Previous analyses showed that the appropriate functional form was linear. All analyses were weighted.

[a] Reported are the marginal means from the multilevel model.

[b] To facilitate comparison of these effect sizes with those reported for the overall VPK effect, effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the standard deviation of the treatment (TN-VPK) group (see Table 3).

[c] For the two-way interactions these are the *p*-values for the interaction terms for ethnicity by treatment condition and primary language by treatment condition. For the three-way interactions, these are the *p*-values for the interaction terms for ethnicity by primary language by treatment condition. Estimates were based on a multilevel model with children nested within pre-k classrooms. In addition to treatment condition and days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) whether the child was Male; (3) whether the child was Black; (4) whether the child was Hispanic; (5) whether the child's native language was not English; (6) whether the child had an IEP placement in the pre-k year; (7) the number of days elapsed between the Woodcock-Johnson testing date and the start of school (centered at the grand mean); and (8) the appropriate interaction term(s). Appendix Tables H-1, I-1 to I-7, J-1, K-1 to K-7, L-1, and M-1 to M-7 present the results for the full models.

As Table 4 reports, the effects of TN-VPK were larger for Hispanic children and for children whose primary language was not English, recognizing that there is considerable overlap in these two variables. Across all bandwidths, the effect sizes for Hispanic children were about double those of their non-Hispanic counterparts. A similar contrast was found for children whose native language was not English compared to those for whom English was their primary language.

The tables in Appendices H through K report detailed results of the analyses for differential effects for Hispanic children and the overlapping native language variable on the subtests of the Woodcock-Johnson measures at the three bandwidths. For the 12- and 6-month bandwidth samples, Hispanic children benefited significantly more than non-Hispanic participants on all of the subtests except Spelling and Oral Comprehension[11] although their mean performance levels were lower than those for their non-Hispanic counterparts in both the No Pre-K and TN-VPK groups (Appendix Table H-1). The effects for children for whom English was not their native language compared with those of native English speakers showed a similar pattern of results with one exception. For the two widest bandwidths, children whose primary language was not English benefited significantly more than their control counterparts on all subtests (including Spelling) except for Oral Comprehension.

For the 3-month bandwidth analyses, however, significant differential effects for Hispanic and non-native English speakers appeared only for Picture Vocabulary and Applied Problems (see Appendix Tables H-1 and J-1). It should be noted that this more restricted sample included relatively small numbers of children who were Hispanic or whose primary language was not English, so statistical power for tests of these interactions was not strong.

Because the majority (76%) of children whose primary language was not English were also Hispanic, there is a question about how much these two comparisons overlap; for example, is the greater TN-VPK effect reported for Hispanic children concentrated in children whose primary language is not English and most likely Spanish? To address this question, the multilevel models were repeated but with the additional three-way interactions that differentiated Hispanic and non-Hispanic children according to their native language. These interaction terms were statistically significant for the 12-month and 6-month bandwidth analyses, though not for the 3-month one (reported in the lower rows of Table 4 for each bandwidth). The patterns across the marginal means show that, although the effects of TN-VPK were positive and large in magnitude for all subgroups, the smallest impact appeared for children who were not Hispanic and whose primary language was English. The subgroup that experienced the largest impact was children whose primary language was not English but who were not Hispanic. These children comprised a quite heterogeneous group that included multiple races and languages (e.g., Arabic, Chinese, Kurdish, Somali, and Vietnamese). Although this pattern characterized all three bandwidth samples, the lack of statistical significance for the 3-month bandwidth may be largely attributable to the small number of children in these differentiated subgroups for that analysis.[12]

Estimated Effects for Non-FRPL Eligible Children

As previously mentioned, children who were not eligible for the federal free or reduced price lunch program (FRPL) could be enrolled in TN-VPK, provided there were available slots and they qualified on

[11] For the 12-month bandwidth, the interaction term was only marginally significant for Oral Comprehension.

[12] Given the small sample sizes in the 3-month bandwidth, analyses for the individual subtests focused on the two widest bandwidth samples and are reported in Appendix Table L-1. Appendix Tables M-1 to M-7 present the estimates for the accompanying models.

other grounds as being at-risk. Narrowing the analytic sample to FRPL- eligible children departs from some previous RD studies of pre-k programs in which all children were included regardless of FRPL status, with this variable simply used as a covariate in the analyses. Regarding the question of whether differential effects exist for children who do and do not qualify for a lunch subsidy, the results have been mixed. For example, Gormley et al. (2005) found significant positive effects on the Letter-Word and Spelling subtests for children regardless of their free lunch eligibility status; for Applied Problems, the results were significant only for children who qualified for free lunch. For the Boston program, however, Weiland and Yoshikawa (2013) found that impacts did not vary on Letter-Word between FRPL-eligible and non-eligible children, but eligible children did benefit significantly more than their ineligible counterparts on the Applied Problems subtest. Thus, focusing on only FRPL-eligible children in the analyses reported above may have limited the generalizability of the results to some degree. While FRPL-eligible children receive priority for admission by many publicly funded programs, they are not the only children who participate.

Based on the limited demographic information available, there were differences between FRPL-eligible and non-eligible children for all three bandwidth samples (see Table 5). Children who qualified for FRPL in the 12-month bandwidth sample were significantly more likely to be Black or Hispanic and to live in a household where the primary language was other than English. Slightly less time also had elapsed from the beginning of the school year to the time that they were administered the WJ subtests, but the difference was reasonably small. Similar nonequivalence appeared in the 6- and 3-month bandwidth samples, again regarding the racial and ethnic composition of the samples along with small differences in the timing of test administration. FRPL-eligible children in both samples were more likely to be Black or Hispanic than non-eligible children, although these differences reached statistical significance only in the 6-month sample, and the time between the beginning of school and test administration was shorter for them. In addition, FRPL-eligible children in the 3-month sample were, on average, younger, but only by less than a week. As indicated by the effect sizes reported in Table 5, most of these differences were relatively small.

Given these differences, analyses were conducted to determine whether there was a different pattern of results when all children were included in contrast to the results reported above for only FRPL-eligible children. Multilevel models similar to those reported in Table 3 were first estimated using the full sample of children but with eligibility status for free or reduced price lunch as a covariate. A second set of models was then estimated that added the interaction of free or reduced price lunch eligibility and treatment condition.

Table 6 summarizes the results from the first set of analyses.[13] Overall, these results were quite similar to those obtained for the sample restricted to FRPL-eligible children. Statistically significant differences favoring the TN-VPK children were found regardless of bandwidth or outcome measure, with one exception (Picture Vocabulary) where the difference was only marginally significant in both 3-month bandwidth samples. The effect sizes also remained quite similar. The average absolute value of the differences between the effect sizes for all children and those for only FRPL-eligible children were small—.03, .03, and .06 for the 12-, 6-month, and 3-month bandwidths, respectively.

[13] Appendix Tables N-1 to N-7 report the results of the full models.

Table 5
Comparison of Baseline Characteristics by Eligibility for Free or Reduced Price Lunch: 12-, 6-, and 3-Month Bandwidth Samples

Variables and Sample	Means[a]		Odds Ratio	TN-VPK vs. No Pre-K Difference (Pooled SD)	Effect Size[b]	*p*-value
	FRPL Eligible	Not Eligible				
12-month bandwidth sample	(N=4144)	(N=1044)				
Gender (Black)	.494	.499	.98		-.01	.785
White	.800	.867	.61		-.27	.002
Black	.204	.118	1.92		.36	<.001
Hispanic	.050	.029	1.75		.31	.002
Primary language is not English	.031	.019	1.67		.28	.016
IEP	.088	.097	.90		-.06	.370
Age at testing (months)	59.4	59.8		-.39 (6.96)	-.06	.104
Days between school start and testing	33.9	34.7		-.72 (5.99)	-.12	.002
6-month bandwidth sample	(N=2067)	(N=535)				
Gender (Black)	.484	.481	1.02		.01	.872
White	.800	.848	.72		-.18	.095
Black	.208	.143	1.57		.25	.024
Hispanic	.052	.033	1.63		.27	.047
Primary language is not English	.040	.026	1.56		.25	.560
IEP	.087	.103	.83		-.10	.254
Age at testing (months)	59.3	59.5		-.26 (3.45)	-.08	.109
Days between school start and testing	33.9	34.6		-.69 (5.81)	-.12	.031
3-month bandwidth sample	(N=1050)	(N=284)				
Gender (Black)	.489	.450	1.17		.09	.248
White	.756	.844	.57		-.31	.025
Black	.251	.154	1.84		.34	.014
Hispanic	.059	.038	1.56		.25	.162
Primary language is not English	.047	.031	1.57		.25	.204
IEP	.090	.118	.74		-.17	.168
Age at testing (months)	59.3	59.6		-.25 (1.78)	.14	.047
Days between school start and testing	33.9	34.7		-.85 (5.78)	-.15	.059

Note. Children regardless of their free or reduced lunch eligibility were included in the analysis. Analyses were weighted.

[a]Reported are estimated means from the multilevel models with children nested within schools, eligibility for free or reduced price lunch as the only predictor, and the respective baseline variable as the dependent variable.

[b]For binary variables (gender, white, black, Hispanic, primary language, and IEP), odds ratios were converted into effect sizes, using Chin's (2000) method. For continuous variables (age and days elapsed since start of school and testing), effect sizes were calculated by dividing the Eligible vs. Not Eligible difference by the pooled standard deviation.

Table 6
Estimates of the Pre-K Effect for the WJ Composite and Subtests: Sample of All Children Regardless of Eligibility for Free or Reduced Price Lunch

Outcome & Sample	Means[a]		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K Difference[b]	*p*	Effect size[c]
	TN-VPK	No Pre-K				
Composite of 6 scales						
± 12 months	412.9	402.0	13.19	10.90	.001	.83
± 6 months	413.2	402.4	12.93	10.82	.001	.84
± 3 months	412.9	402.6	12.21	10.30	.001	.84
Letter Word						
± 12 months	347.8	326.2	21.79	21.66	.001	.99
± 6 months	348.0	326.8	20.69	21.25	.001	1.03
± 3 months	346.7	327.6	20.14	19.08	.001	.95
Spelling						
± 12 months	380.6	360.3	21.06	20.35	.001	.97
± 6 months	381.3	361.3	20.76	20.01	.001	.96
± 3 months	380.8	361.0	20.34	19.87	.001	.98
Oral Comprehension						
± 12 months	452.8	448.9	13.88	3.86	.001	.28
± 6 months	452.8	449.3	13.64	3.50	.001	.26
± 3 months	453.3	449.6	13.28	3.71	.009	.28
Picture Vocabulary						
± 12 months	465.3	461.9	10.73	3.40	.001	.32
± 6 months	465.4	462.2	11.00	3.26	.002	.30
± 3 months	465.4	462.8	10.36	2.65	.063	.26
Applied Problems						
± 12 months	409.5	401.0	18.01	8.53	.001	.47
± 6 months	411.0	401.1	17.98	9.93	.001	.55
± 3 months	410.6	401.0	16.16	9.52	.001	.59
Quantitative Concepts						
± 12 months	421.5	413.9	14.40	7.55	.001	.52
± 6 months	420.9	413.9	13.92	7.02	.001	.50
± 3 months	420.5	413.7	13.25	6.83	.001	.52

Note. The sample sizes for the TN-VPK and No Pre-K groups were: 2,621 and 2,567 for ±12 months; 1,280 and 1,322 for ±6 months; and 661 and 673 for ±3 months. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Full results of the multilevel models are reported in Appendix Tables N-1 to N-7.

[a] Estimated marginal means.

[b] Estimates based on a multilevel model with children nested within pre-k classroom. In addition to days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) Male; (3) Black; (4) Hispanic; (5) non-native English; (6) IEP placement; (7) FRPL eligible; and (8) days between the WJ testing and the start of school (grand mean centered).

[c] Effect sizes are the TN-VPK vs. No Pre-K difference divided by the standard deviation of the TN-VPK treatment group.

The second set of analyses included interaction terms to explicitly compare the effects of TN-VPK for FRPL-eligible children versus non-eligible children. The results for the WJ composite and the WJ subtests with each bandwidth are reported in Table 7. For the 12-month bandwidth, the estimated benefit of TN-VPK was significantly larger for FRPL-eligible children on the WJ composite and four of the six Woodcock-Johnson subtests (Letter-Word, Picture Vocabulary, Applied Problems, and Quantitative Concepts). However, no differences between the FRPL groups reached statistical significance in the analyses with the narrower bandwidths.[14]

[14] Appendix Tables O-1 to O-7 report the results for the full models.

Table 7
TN-VPK Effect Estimates for WJ Scores by Children's Free or Reduced Lunch Eligibility

Outcome and Sample	TN-VPK		No Pre-K		Pooled SD	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean	*N*	Mean	*N*				
Composite								
12-month bandwidth					18.93			.005
Eligible	412.3	2078	400.8	2066		11.52	.61	
Not eligible	415.6	543	406.7	501		8.87	.47	
6-month bandwidth					16.74			.358
Eligible	412.1	1009	401.0	1058		11.12	.66	
Not eligible	417.6	271	407.6	264		9.96	.60	
3 -month bandwidth					16.08			.644
Eligible	411.4	519	401.3	531		10.11	.63	
Not eligible	418.3	142	407.3	142		10.91	.68	
Letter-Word								
12-month bandwidth					29.32			
Eligible	346.8	2078	324.3	2066		22.52	.77	.015
Not eligible	352.2	543	333.3	501		18.82	.64	
6-month bandwidth					27.15			.406
Eligible	346.3	1009	324.6	1058		21.67	.80	
Not eligible	354.9	271	335.0	264		19.92	.73	
3 -month bandwidth					26.54			.834
Eligible	344.8	419	325.5	531		19.24	.73	
Not eligible	353.9	142	335.3	142		18.60	.70	
Spelling								
12-month bandwidth					29.96			.428
Eligible	379.7	2078	359.1	2066		20.60	.69	
Not eligible	384.5	543	365.0	501		19.49	.65	
6-month bandwidth					26.45			.708
Eligible	379.9	1009	359.7	1058		20.18	.76	
Not eligible	386.6	271	367.2	264		19.43	.74	
3 -month bandwidth					25.24			.253
Eligible	378.7	519	359.6	531		19.10	.76	
Not eligible	388.6	142	366.2	142		22.36	.89	
Oral Comprehension								
12-month bandwidth					15.67			
Eligible	452.2	2078	448.3	2066		3.89	.25	.887
Not eligible	455.1	543	451.4	501		3.77	.24	
6-month bandwidth					14.57			.505
Eligible	451.8	1009	448.5	1058		3.31	.23	
Not eligible	456.6	271	452.5	264		4.12	.28	
3 -month bandwidth					14.32			.092
Eligible	452.0	419	448.9	531		3.04	.21	
Not eligible	458.1	142	452.2	142		5.90	.41	

Table 7 (*continued*)
TN-VPK Effect Estimates for WJ Scores by Children's Free or Reduced Lunch Eligibility

Outcome and bandwidth	TN-VPK		No Pre-K		Pooled SD	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Picture Vocabulary								
12-month bandwidth					15.21			.004
Eligible	465.1	2078	461.1	2066		3.98	.26	
Not eligible	466.3	543	464.8	501		1.51	.10	
6-month bandwidth					14.81			.225
Eligible	465.0	1009	461.4	1058		3.61	.24	
Not eligible	467.3	271	465.2	264		2.13	.14	
3 -month bandwidth					14.92			.421
Eligible	464.8	519	461.8	531		2.97	.20	
Not eligible	467.8	142	466.2	142		1.60	.11	
Applied Problems								
12-month bandwidth					26.04			
Eligible	409.2	2078	399.3	2066		9.97	.38	<.001
Not eligible	411.2	543	407.4	501		3.81	.15	
6-month bandwidth					22.86			.215
Eligible	410.0	1009	399.5	1058		10.48	.46	
Not eligible	415.1	271	406.9	264		8.14	.36	
3 -month bandwidth					22.29			.747
Eligible	409.0	519	399.7	531		9.32	.42	
Not eligible	416.3	142	406.1	142		10.15	.46	
Quantitative Concepts								
12-month bandwidth					17.27			.007
Eligible	420.9	2078	412.8	2066		8.11	.47	
Not eligible	423.9	543	418.1	501		5.73	.33	
6-month bandwidth					15.46			.296
Eligible	420.0	1009	412.6	1058		7.33	.47	
Not eligible	424.7	271	418.7	264		6.01	.39	
3 -month bandwidth					14.91			.956
Eligible	419.4	519	412.6	531		6.80	.46	
Not eligible	424.8	142	417.9	142		6.90	.46	

Note. Children regardless of their free or reduced lunch eligibility were included in the analysis. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

[a] Reported are estimated marginal means from the multilevel analysis model.

[b] Effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the pooled standard deviation.

[c] *p*-value for the interaction term for condition (TN-VPK vs. No Pre-K) and FRPL eligibility. Estimates were based on a multilevel model with children nested within pre-k classroom. In addition to days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) Male; (3) Black; (4) Hispanic; (5) non-native English; (6) IEP placement; (7) FRPL eligible; and (8) days between the WJ testing and the start of school (grand mean centered); and (9) the interaction term. Appendix Tables O-1 to O-7 report the full results for each model.

Summary and Conclusions

As noted earlier, the age-cutoff regression-discontinuity design has been used in a number of studies of publicly-funded pre-k programs. Like the present study, most of these have focused on children's cognitive skills, especially emergent literacy, language, and mathematics abilities, with a few also including measures of social, emotional, or behavioral outcomes. While a great strength of the RD design is the relative ease of implementing it as a practical matter in pre-k settings, it does have some limitations. Most notably, the inherent nature of the age-cutoff RDD as essentially a wait-list control design limits assessment of pre-k effects to those appearing shortly after the end of the pre-k year, so the question of sustained effects cannot be investigated.

Further, although regression-discontinuity designs in general are capable of yielding unbiased causal impact estimates, a host of methodological criteria must be satisfied for this to be accomplished. These include an exogenously imposed cutoff, application of statistical models that correctly depict the functional form of the relationship between the cutting-point variable and the outcomes, and the choice of an appropriate bandwidth around the cutting point.

In addition, the pre-k age-cutoff version of the RDD has some distinctive characteristics that impose additional demands (Lipsey, Weiland, et al., 2015). These stem primarily from the inherent comparison of two discrete cohorts separated by time as well as age when an age-eligibility cutoff is used to define the treatment and control groups. This circumstance opens the door to cohort differences that may bias the effect estimates, most obviously the possibility of trends that change the demographic mix of the children who enroll in pre-k in successive years. Methodological differences may also be in play, for example, for the consistency with which outcome measures are operationalized and administered across the two cohorts, one tested in pre-k, the other in kindergarten, and the possibility of differential attrition in obtaining those measures.

In our RDD evaluation of the TN-VPK program, we attempted to address these issues to the extent possible given the practical limitations of field-based research. Meeting the widely recognized formal criteria for regression-discontinuity designs was relatively straightforward. The strict age cutoff used for VPK eligibility each school year provided a well-defined and exogenously imposed cutpoint for differentiating pre-k participants from analogous nonparticipants. A thorough exploration of the functional form of the relationship between age and the different outcome measures was conducted and rather conclusively identified a simple linear model as a good fit for all those outcomes. The effect estimates were repeated with different bandwidths around the cutpoint to assess their robustness and did not show enough variation to call the overall findings into question.

More challenging was the need to ensure the equivalence of the children compared across the cohorts on characteristics potentially related to the outcomes of interest. We were limited to the small set of baseline demographic variables available in the state data system for children participating in VPK but included those as covariates in all the analytic models to statistically adjust for any differences between cohorts on those variables. In addition, we tried to maximize the comparability of the TN-VPK and No Pre-K cohorts by imposing identical study eligibility criteria. Children were only included in the analytic sample if they were in the same VPK classrooms with the same teacher the year before for the VPK cohort and the beginning of the current year for the No Pre-K cohort, save for a few exceptions that arose for practical reasons. Moreover, the children in both cohorts had to have been enrolled within the beginning weeks of the appropriate pre-k year, remain enrolled through nearly the end of the school year, and

attend a Tennessee public school for kindergarten after their respective pre-k year. Also, efforts were made to make the administration of the outcome measures as systematic and consistent across cohorts as possible and the amount of elapsed time between the start of the school year and testing was used as a covariate in the analyses.

While we believe these study features support the internal validity of the pre-k effect estimates obtained, the question of their external validity is also relevant. In that regard, this study is rather distinctive among pre-k age-cutoff RDDs for its use of a statewide probability sample of TN-VPK classrooms. That aspect allows relatively confident generalization of the results to the entire Tennessee program, at least as it was implemented in the years over which the study data were collected. The extent to which other state programs produce comparable effects is an open question, but we would note that the general characteristics of TN-VPK are quite similar to those of many of those programs.

Summary of Findings

The main focus of the analysis of TN-VPK effects in this study was on children who qualified for the federal free or reduced price lunch programs, an indicator of low family income. By statute, that group is given the highest priority for admission to VPK and, indeed, the great majority of children in the program meet this criterion. The effect size estimates for these children on the WJ Composite outcome measure that averaged across all the separately assessed subtests ranged from .85 to .91 for analyses with different bandwidths. Effect sizes are in standard deviation units and can be easily translated into percentile differences when the outcome data are normally distributed, as these are. The mean of the No Pre-K control group score in this translation would be at the 50th percentile on the distribution of their scores. A difference of .85 standard deviations between the mean for this control group and that for the TN-VPK treatment group would place the mean for TN-VPK participants at the 80th percentile on that distribution. This represents a rather considerable relative improvement in the measured cognitive skills of the children who participated in TN-VPK programs.

The estimated TN-VPK effects, however, varied quite a bit across the subtests measuring literacy, language, and mathematics skills. The largest effects were found for the literacy measures (WJ Letter Word and Spelling tests). Those effect sizes ranged from .94 to 1.11 across these two subtests and the different bandwidths. The smallest effects by a substantial margin were found on the language measures (WJ Oral Comprehension and Picture Vocabulary). The effect sizes for those outcome measures ranged from .22 to .34 across the respective bandwidth analyses. The effect sizes for the mathematics measures (WJ Applied Problems and Quantitative Concepts) fell in between, ranging from .47 to .71 across these two subtests and the different bandwidths.

Subgroup differences on VPK effects were examined for the few characteristics of the children in the sample that were documented in the available data—gender, race/ethnicity, and primary language. The results revealed that VPK effects on the overall WJ Composite measure were significantly larger for Hispanic children in contrast to non-Hispanic children, and for children whose primary language was not English in contrast to those with English as their native language. These are clearly overlapping characteristics. Further analysis showed that the largest VPK effects were found for the relatively small subgroup of children who were not Hispanic but also did not have English as their primary language. The smallest, but still positive, effects were found for non-Hispanic native English speaking children. VPK effects on this overall composite achievement measure for the Hispanic children whose primary

language was English and those whose primary language was not English fell in the middle range and were similar for these two subgroups.

While the focus of this study was on children who qualified for free or reduced price lunch and were thus the priority target group for TN-VPK, analysis was also conducted for the full sample including both FRPL and non-FRPL eligible children. That more inclusive sample included about 25% more children than the FRPL-only sample. Nonetheless, the pattern of results for that more inclusive sample was substantially similar to that for the FRPL-only sample. The effect sizes for the WJ Composite and WJ subtests were within ±.05 on the 12-month bandwidth estimates for the two samples with those for the more inclusive sample tending to be slightly lower when there was a difference.

A direct test of the differences between the FRPL-only sample and the non-FRPL sample showed that indeed there were larger VPK effects for the FRPL-only sample that were statistically significant in the 12-month bandwidth analyses for the WJ Composite and the Letter Word, Picture Vocabulary, Applied Problems, and Quantitative Concepts subtests. Although the pattern of differences on these outcome measures was similar for the narrower bandwidth analyses, none of those reached statistical significance.

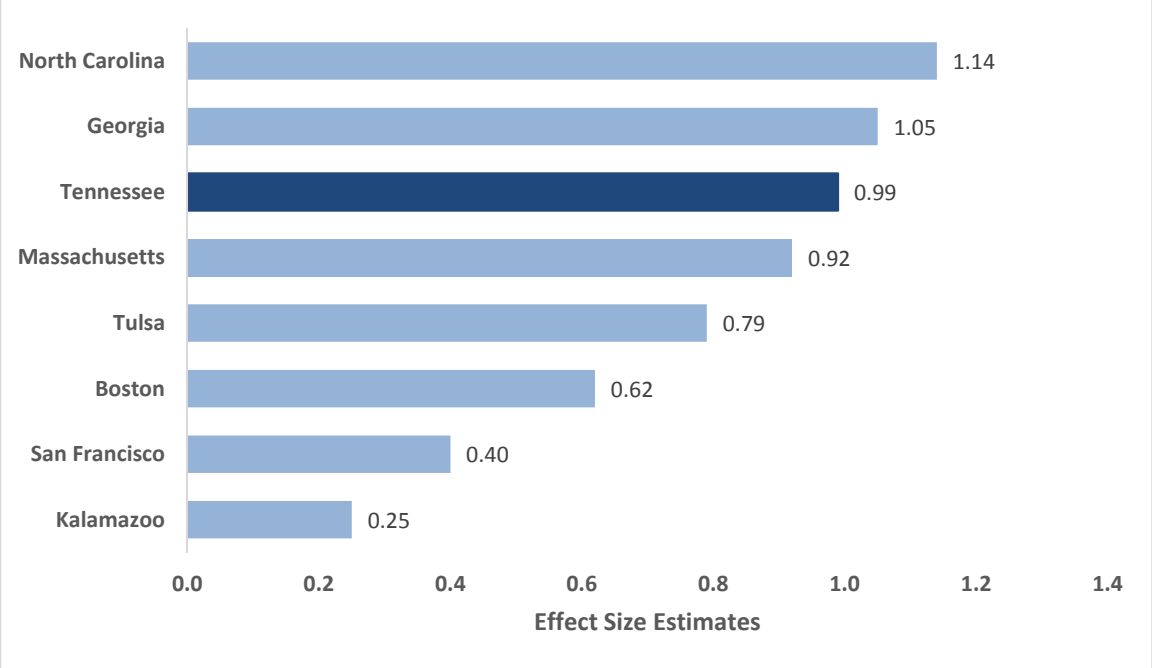
Taken together, the subgroup analyses indicate that the effects of VPK were somewhat larger for the children from low-income families eligible for the FRPL programs and notably larger for Hispanic children and the overlapping group of children for whom English was not their native language, whether Hispanic or not.

Comparison with the effects found in other pre-k age-cutoff RD studies. The extent of the gains made by children who attended VPK relative to those who had not yet attended is one index of the effectiveness of TN-VPK. In that regard, the positive statistical effect sizes found in this study, especially the rather large ones for the literacy and math measures, demonstrate that TN-VPK has achieved some success in improving the readiness of the low-income students it serves for entry into kindergarten.

One way to assess that success is to compare the magnitude of the VPK effects with those found in comparable studies of other state or locally funded pre-k programs. We have identified studies of 18 other such pre-k programs that used the age-cutoff regression-discontinuity design to estimate program effects at the beginning of the kindergarten year (details in Appendices P-1 and P-2). These studies vary in how carefully the RDD was implemented and analyzed and in the nature of the programs and the samples of children used in the respective studies. Nonetheless, it is informative to examine the distribution of effect sizes found across these similarly designed studies and, in the present context, to observe where the TN-VPK effects fall in that distribution.

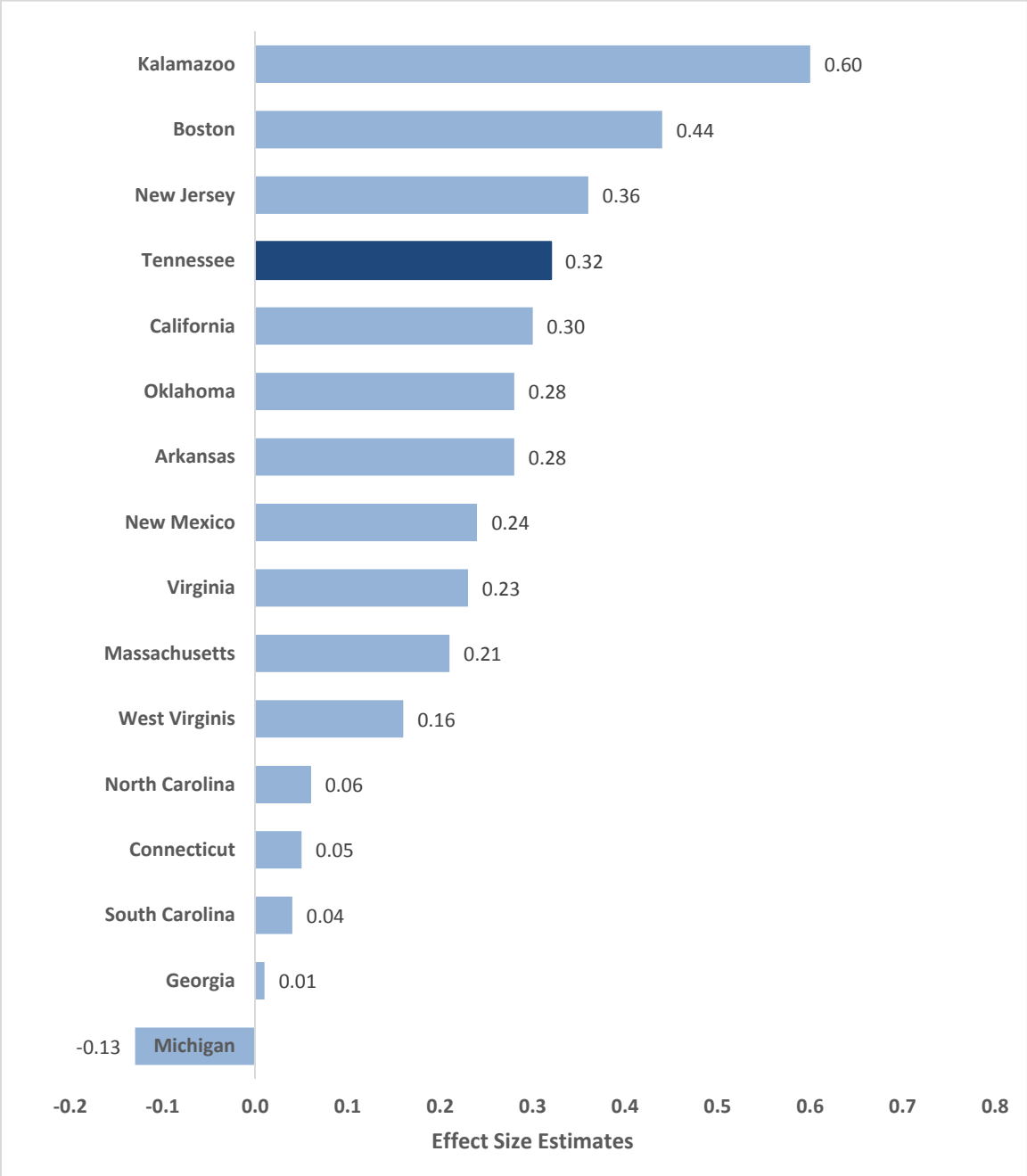
Overall, these RDD studies show overwhelmingly positive effects for the pre-k programs studied, especially in the almost universally measured outcome domains of literacy, language, and math skills. Fortunately for purposes of comparison, many of these studies use the same or very similar measures of these skills. Literacy in particular is frequently measured with the WJ Letter-Word Identification subtest, language with either the Peabody Picture Vocabulary Test (PPVT) or the WJ Picture Vocabulary subtest, and math with the WJ Applied Problems subtest. These measures have also been used in this study of the TN-VPK program, allowing the effects found to be rather directly compared with those reported in many of these other studies. Figure 3 shows that comparison for WJ Letter-Word Identification, Figure 4 for PPVT or WJ Picture Vocabulary, and Figure 5 for WJ Applied Problems.

Figure 3
Effect Size Estimates for WJ Letter-Word Identification from Age-Cutoff Regression-Discontinuity Studies of Pre-K Programs



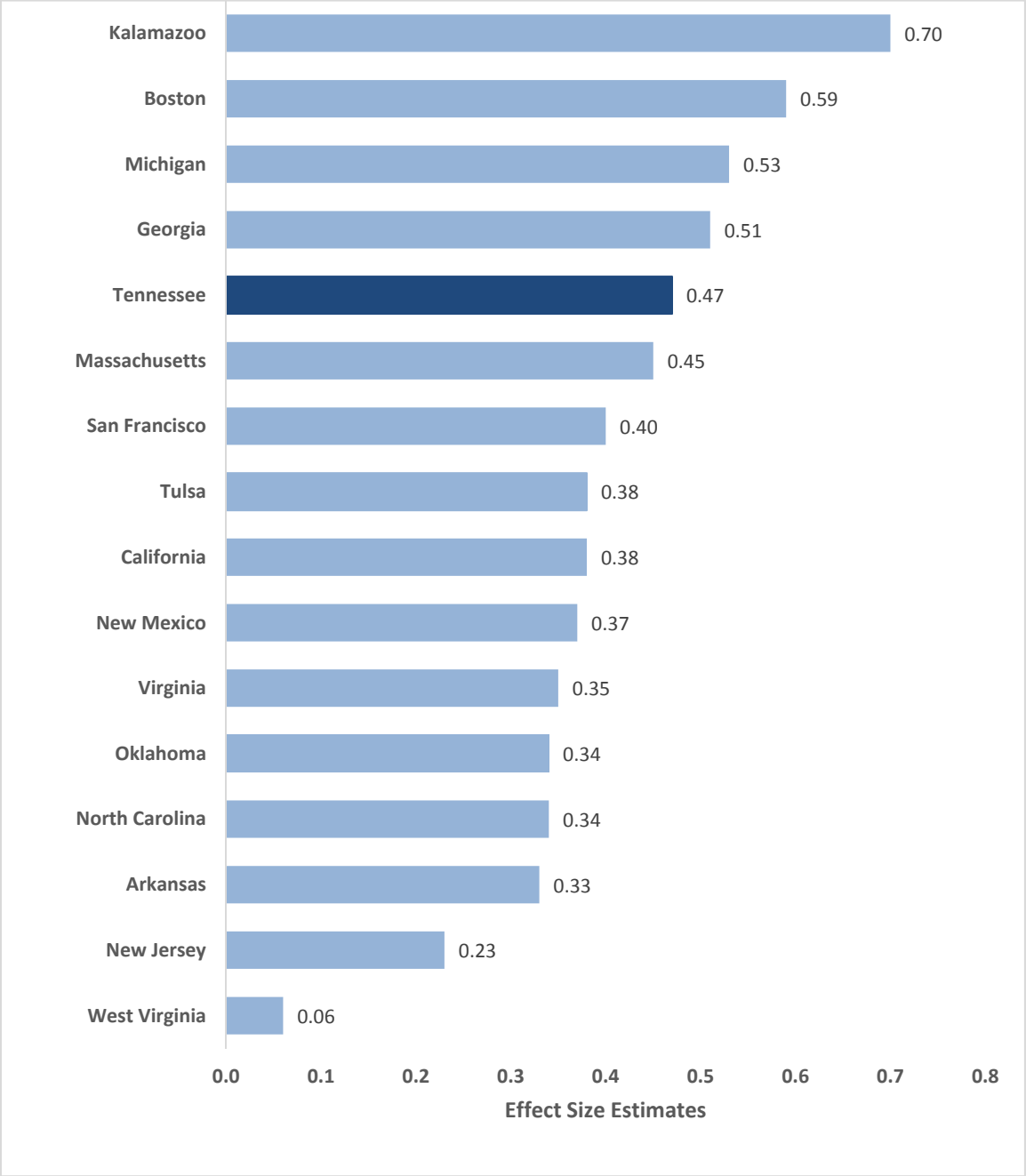
Note. Appendix P-1 identifies the reported effect size estimates for all outcomes examined in each study. References for the source studies are in Appendix P-2.

Figure 4
Effect Size Estimates for PPVT or WJ Picture Vocabulary from Age-Cutoff Regression-Discontinuity Studies of Pre-K Programs



Note. Appendix P-1 identifies the reported effect size estimates for all outcomes examined in each study. References for the source studies are in Appendix P-2.

Figure 5
Effect Size Estimates for WJ Applied Problems from Age-Cutoff Regression-Discontinuity Studies of Pre-K Programs



Note. Appendix P-1 identifies the reported effect size estimates for all outcomes examined in each study. References for the source studies are in Appendix P-2.

Figures 3-5 reveal the generally positive effects on these outcome measures found for the state and locally funded pre-k programs studied with the age-cutoff RDD. The TN-VPK effect sizes shown in these distributions are for the full sample, not the effect sizes that were generally somewhat larger for the FRPL-eligible sample that was the primary focus of this study. Not all the programs represented in these distributions prioritize FRPL-eligible children the way TN-VPK does and, indeed, some of them are universal programs open to virtually all age-eligible children.

While acknowledging the methodological and programmatic variation across these studies that compromise any direct comparison, it is encouraging for the Tennessee Department of Education that the TN-VPK effects compare relatively well with those from these other programs. While TN-VPK is not at the very top of any of these distributions, it is in the top half in all three, showing what can be described as an above average performance compared to the peer programs for which RDD data are available. However, it is important to recognize the differences in the order of magnitude of the effect sizes across the literacy, language, and math domains. The overall pattern in all these studies, which is consistent with the VPK findings, shows the largest effects on literacy skills and the smallest on language skills, with math in between. While this is understandable given the emphasis on emergent literacy typical of pre-k programs, it does raise questions about whether early language and math are receiving sufficient attention to build the skills participating children need to be well prepared for kindergarten and the subsequent grades.

Comparison with the effects found in the parallel RCT. As mentioned early in this report, the RDD was one component of a larger TN-VPK study that also included a randomized control component, referred to as the RCT, which compared outcomes for applicants offered admission to oversubscribed TN-VPK programs with those waitlisted and ultimately not offered admission. One part of that RCT, which we called the intensive substudy (ISS), used the same Woodcock Johnson achievement measures as the RDD with a sample of consented children tested at the beginning and end of the pre-k year. Details for the RCT and ISS, and the associated findings, are reported elsewhere (Lipsey, Farran, & Hofer, 2015; Lipsey, Farran, & Durkin, 2018).

As a consented subsample of the RCT, the ISS was analyzed as a quasi-experiment, but had the advantage of an extensive set of baseline measures used to adjust for any initial differences on those measures between the treatment and control groups. While the design and sample were quite different in the ISS and the RDD, both generated estimates of TN-VPK effects on the selected WJ measures. The effect sizes on each of the WJ measures they had in common are compared in Table 8. Because the RDD samples represent children who actually participated in VPK, the ISS effect estimates in Table 8 are those from the treatment-on-the-treated (TOT) analysis that also defines the treatment and control groups in terms of actual participation irrespective of the condition to which the children were originally assigned in the RCT.

Table 8 shows some similarity in the overall pattern of the effect estimates. Most important, perhaps, both the RDD and ISS found positive VPK effects on all the measures, although that for Oral Comprehension in the ISS is quite small and not statistically significant. Further, both the RDD and ISS show a generally similar pattern of effects across the various outcomes. The effect estimates are largest for the literacy outcomes and those for the language outcomes are among the smallest, with the math outcomes in between. However, these differences are not as distinct for the ISS effect sizes. There is thus some mutual confirmation of the generally positive VPK effects and their pattern across outcome measures between these two substudies.

Table 8
Effect Sizes from the RDD Compared with Those from the ISS

Outcome Measure	Effect Size Estimates	
	RDD[a]	ISS[b]
WJ Composite	.85	.40
Literacy		
Letter-Word Identification	1.04	.47
Spelling	.98	.38
Language		
Oral Comprehension	.24	.07
Picture Vocabulary	.34	.32
Mathematics		
Applied Problems	.47	.26
Quantitative Concepts	.56	.33

Notes: Estimates for both the RDD and ISS are from samples of FRPL eligible children. All these effect sizes are statistically significant except for ISS Oral Comprehension.

[a] 12-month bandwidth estimates.

[b] Treatment-on-the-treated estimates.

At the same time, there is a quite noticeable difference between the RDD and the ISS in the magnitude of the effect estimates. The RDD effect sizes are larger than those from the ISS in every instance, with that difference quite large for some measures. For the overall WJ Composite measure, as well as some subtests, the RDD effect sizes are more than twice as large as the ISS estimates. It is difficult to interpret those effect size differences given the many differences between the sources of the two sets of estimates. Most notably, the RDD used a probability sample of statewide TN-VPK programs whereas the ISS used a consented subsample of the RCT sample of selected oversubscribed programs. Also, outcome assessment in the RDD occurred at the beginning of the kindergarten year, approximately 12 months after initial pre-k enrollment, but at the end of the pre-k school year in the ISS, approximately 9 months after initial pre-k enrollment. Moreover, outcome measurement for the control group in the RDD came at the beginning of the pre-k year, making the prior 3-year old period the counterfactual condition. In contrast, outcome measurement for the ISS control group at the end of the pre-k school year made the 4-year old period the counterfactual condition for those children. The implications of these and other relevant differences between the RDD and ISS for the magnitude of the respective effect size estimates are being explored in a separate report.

Conclusions

What is most clearly demonstrated in this RDD study is that the TN-VPK program, viewed statewide, has positive effects on the cognitive skills that should make the participating children better prepared for the learning opportunities they will encounter in kindergarten. The gains made by the participating children, however, are somewhat unevenly distributed. Those gains are much stronger for early literacy skills than for the early language skills that arguably may be equally or more important to later academic progress. And, similarly, the gains in early math skills, while positive, lag those in literacy.

The distribution of VPK gains across the major demographic subgroups of children, on the other hand, is rather even. While positive effects were found across all these subgroups, the effects were larger for Hispanic children and children for whom English was not their primary language, whether Hispanic or not. Otherwise, however, VPK was equally effective for the other race/ethnic groups and for boys and girls. The separate analysis of differential effects associated with FRPL eligibility, did show somewhat larger gains for the eligible children, thus supporting the TN-VPK focus on FRPL eligible children as a priority.

REFERENCES

Applied Survey Research (2013). *Evaluating preschool for all effectiveness*. Research Brief. San Jose CA: Applied Survey Research.

Bartik, T. J. (2013). *Effects of the pre-k program of Kalamazoo County Ready 4s on kindergarten entry test scores: Estimates based on data from the fall of 2011 and the fall of 2012*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, Upjohn Institute Working Paper 13-198.

Chin, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, 19, 3127-3131.

Doss, C. (2016). *Transitional kindergarten vs. Prekindergarten: A fuzzy regression discontinuity analysis of student literacy skills* (CEPA Working Paper No. 16-07). Palo Alto, CA: Stanford Center for Education Policy Analysis, Stanford University.

Frede, E. (2005). *Assessment in a continuous improvement cycle: New Jersey's Abbott Preschool Program*. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.

Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-k on cognitive development. *Developmental Psychology*, *41(6)*, 872-884.

Huang, F. L. (2017). Does attending a state-funded preschool program improve letter name knowledge? *Early Childhood Research Quarterly*, 30, 116-126.

Hustedt, J. T., Barnett, W. S., Jung, K., & Thomas, J. (2007). *The effects of the Arkansas Better Chance Program on young children's school readiness*. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.

Hustedt, J. T., Barnett, W. S., Jung, K., Goetze, L. D. (2009). *The New Mexico PreK evaluation: Results from the initial four years of a new state preschool initiative—Final report*. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.

Jacob, R., Zhu, P. Somers, M-A., & Bloom, H. (2012). *A practical guide to regression discontinuity*. New York, NY: MDRC.

Jaffe, L. E. (2009). *Development, interpretation, and application of the W score and the relative proficiency index* (Woodcock-Johnson III Assessment Service Bulletin No. 11). Rolling Meadows, IL: Riverside Publishing.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48(2)*, 281-355.

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, *45*, 155-176. https://reader.elsevier.com/reader/sd/pii/S0885200618300279?token=A7204D8451402AA722801F532F80B15A23C38850C266544B3CD8906A5536ECC0FC13C2B52E1A07E042E1F6D61CF55DAC

Lipsey, M. W., Farran, D.C., & Hofer, K. G., (2015). *A randomized control trial of the effects of a statewide voluntary prekindergarten program on children's skills and behaviors through third grade* (Research Report). Nashville, TN: Vanderbilt University, Peabody Research Institute. https://my.vanderbilt.edu/tnprekevaluation/technical-reports/

Lipsey, M. W., Hofer, K. G., Dong, N., Farran, D. C., & Bilbrey, C. (2013a). *Evaluation of the Tennessee Voluntary Prekindergarten Program: End of pre-k results from the randomized control design*. Research Report. Nashville, TN: Vanderbilt University, Peabody Research Institute. https://my.vanderbilt.edu/tnprekevaluation/technical-reports/

Lipsey, M. W., Hofer, K. G., Dong, N., Farran, D. C., & Bilbrey, C. (2013b). *Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and first grade follow-up results from the randomized control design*. Research Report. Nashville, TN: Vanderbilt University, Peabody Research Institute. https://my.vanderbilt.edu/tnprekevaluation/technical-reports/

Lipsey, M. W., Weiland, C., Yoshikawa, H., Wilson, S. J., & Hofer, K. G. (2015). The prekindergarten age-cutoff regression-discontinuity design: Methodological issues and implications for application. *Educational Evaluation and Policy Analysis*, *37(3)*, 296-313.

McGrew, K. S, Schrank, K. A., & Woodcock, R. W. (2007). *Woodcock-Johnson III normative update technical manual*. Rolling Meadows, IL: Riverside.

Montrosse-Moorhead, B., La Salle, T., Dougherty, S., Dostal, H., & Weiner, J. (June, 2016). *Early childhood regression-discontinuity study: A report by the Connecticut Academy of Science and Engineering for the Connecticut General Assembly and Education Committee.* Rocky Hill, CT: Connecticut Academy of Science and Engineering.

Peck, L. R., & Bell, S. (2014). *The role of program quality in determining Head Start's impact on child development*. OPRE Report #2014-10. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U. S. Department of Health and Human Resources.

Peisner-Feinberg, E. S., Schaaf, J. M., LaForett, D. R., Hildebrandt, L. M., & Sideris, J. (2014). *Effects of Georgia's Pre-K Program on children's school readiness skills: Findings from the 2012-2013 evaluation study*. Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, D., Mashburn, A., & Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study final report*, OPRE Report # 2012-45. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U. S. Department of Health and Human Resources.

Vogel, C. A., Xue,., Y., Moiduddin, E., Kisker, E., & Carlson, B. L. (2010). *Early Head Start children in grade 5: Long-term follow-up of the Early Head Start research and evaluation project study sample.* OPRE Report #2011-8. Washington, DC: U. S. Department of Health and Human Services, Administration for Children and Families, Office for Planning, Research, and Evaluation.

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84* (6), 2112-2130.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Tests of Cognitive Abilities-III*. Itasca, IL: Riverside Publishing Company.

Appendix A. The RDD Sampling Strategy

A probability sampling strategy was used to construct a sample representative of TN-VPK classrooms in Tennessee. The state was first divided into four geographic regions that closely approximated the three "Grand Divisions" (East, Middle, and West) widely recognized in Tennessee geography, but with the large Middle Division split into two regions (Central West and Central East). The school districts within each of these regions were enumerated and schools with a TN-VPK program were identified within each district. Each TN-VPK program is associated with a single host school but may include more than one classroom. Overall, there were 132 districts that included at least one school with a TN-VPK program for a total of 646 programs with 942 TN-VPK classrooms within those districts and schools (see Appendix Table A-1).

Each TN-VPK program within each region was then assigned to a stratum based on its profile on four characteristics that Tennessee Department of Education personnel identified as important distinctions within the TN-VPK program—urban/nonurban, partnership sites, pilot sites, and AYP priority schools. Urban/nonurban distinguished programs in the major urban areas. Partnership sites were those where the VPK programs were administratively associated with specific schools, but operated by a community center or Head Start agency. Pilot sites were the locations of the original and oldest VPK programs implemented during the initial pilot phase of the TN-VPK roll out. AYP priority schools were those TN-VPK host schools designated as underperforming by the state and thus eligible for a range of supports and accountability monitoring.

As Appendix Table A-1 reports, the combinations of these program characteristics produced 11 distinct strata within regions, with each region having between 6 and 11 strata. One classroom from each program that met the eligibility requirements described below was included in the sampling frame. In three of the strata there was at least one program with no classroom drawn from that stratum. These strata were merged within region with the next most similar stratum, with similarity determined by allowing, first, differences on AYP priority, then on pilot site status.

For a classroom to be eligible for selection into the final sample, it had to be in place for both the 2009-10 and 2010-11 school year and staffed by the same teacher both years. The requirement that both age cohorts of children be enrolled in the same TN-VPK classrooms was imposed to help make the children in the two cohorts as comparable as possible on whatever unmeasured variables were associated with the local school catchment areas and the process by which children were sorted into classrooms.

Within each region, a disproportionate random sample of eligible classrooms was selected with the aim of representing every strata but with larger sampling proportions for the smaller strata and smaller sampling proportions for the larger ones. Sampling weights were generated for use in the analysis that differentially weighted the classrooms in the different strata to make the resulting proportions match those for the statewide population of TN-VPK classrooms. The

original sampling frame included 942 classrooms. Across the four regions, the final sample totaled 155 classrooms from 154 schools (one school was allowed to contribute 2 classrooms) and 73 school districts.

In constructing this final sample, various sampling-related decisions were made in the different regions. These were as follows:

- **West.** The West region consists of 36 school districts and includes the city of Memphis. Forty TN-VPK programs in 16 of the 36 school districts were drawn for the sample. All 40 programs contributed a classroom to the final participating sample in this region.
- **Central West.** The Central West region has 28 middle Tennessee school districts, including the city of Nashville. TN-VPK programs in 17 of the districts in this region were drawn for the sample. Originally, 39 classrooms were selected from those programs, but three were dropped because changes at the host schools prevented the second cohort of children (the control group) from enrolling in TN-VPK classrooms comparable to the ones in which the first cohort (the treatment group) had participated the year before, leaving 36 classrooms in the sample from this region.
- **Central East**. This region includes 32 districts and is the region in which Chattanooga is located. The initial sample had 40 TN-VPK programs in 20 of these districts. However, only 39 classrooms from those programs remained in the final sample; one was excluded due to the lack of comparability between the first and second cohort of children.
- **East.** In the East region, which includes 36 school districts and the city of Knoxville, 20 districts that housed 41 VPK programs were sampled. Forty of the 41 programs contributed a classroom to the final sample; one classroom had to be dropped because it did not have a comparable second cohort of children.

Construction of sampling weights. As noted, the original sampling frame included 942 TN-VPK classrooms, the full population of TN-VPK classrooms located by the research team at the time this study began. Of those, 155 classrooms were selected for the final sample. The sampling fraction for each stratum was computed as the number of classrooms sampled from that stratum divided by the total number of classrooms in that stratum in the original sampling frame. Thus, the overall sampling fraction was 155/942=.1645.

The weight to be assigned to each classroom in the analyses was calculated as the overall sampling fraction (.1645) divided by the sampling fraction for the respective stratum. Multiplying that weight by the number of cases sampled in each stratum yielded the number of cases from the population represented by that stratum. These summed to 155 as expected, i.e., the total sample *N* remained the same with the weighting.

Appendix Table A-1
Sampling Frame for VPK Classrooms, Strata, and Sampling Fractions

Region	Number of Districts		Strata				Number of classrooms		Sampling fraction
	Total	Sample	Urban	Partner-ship	Pilot	AYP Priority	Total	Sampled	
Central West	28	17	X				47	6	0.1277
			X			X	6	3	0.5000
			X		X		6	3	0.5000
			X	X			5	3	0.6000
			X	X	X		4	1	0.2500
						X	2	1	0.5000
					X		47	7	0.1489
				X			3	1	0.3333
							117	11	0.0940
Central West Totals	**28**	**17**					**237**	**36**	
West	36	16	X				47	6	0.1277
			X			X	11	3	0.2727
			X		X		9	3	0.3333
			X	X			39	6	0.1538
			X	X	X		5	2	0.4000
						X	6	2	0.3333
					X		33	3	0.0909
					X	X	15	3	0.2000
				X			14	3	0.2143
				X	X		1	1	1.0000
							101	8	0.0792
West Totals	**36**	**16**					**281**	**40**	

Appendix Table A-1 (*continued*)
Sampling Frame for VPK Classrooms, Strata, and Sampling Fractions

Region	Number of Districts		Strata				Number of classrooms		Sampling fraction
	Total	Sample	Urban	Partner-ship	Pilot	AYP Priority	Total	Sampled	
Central East	32	20	X				9	3	0.3333
			X			X	1	1	1.0000
			X		X		4	1	0.2500
			X	X			13	4	0.3077
						X	4	2	0.5000
					X		16	5	0.3125
				X			8	4	0.5000
							153	19	0.1242
Central East Totals	**32**	**20**					**208**	**39**	
East	36	20	X				7	4	0.5714
						X	3	2	0.6667
					X		31	9	0.2903
				X			18	6	0.3333
				X	X		10	3	0.3000
							147	16	0.1088
East Totals	**36**	**20**					**216**	**40**	
Grand Totals	**132**	**73**					**942**	**155**	**.1645**

Note. The overall sampling fraction was 0.1645, which was divided by the sampling fraction of the respective stratum to derive the overall weight for that stratum. The regions are listed in the order in which data collection occurred. That is, outcome data were collected for Central West schools in the fall of 2010, for West schools in fall 2011, for Central East schools in fall 2012, and for East schools in fall 2013.

Appendix Table B-1
Statistical Models and Results for the WJ Composite

Independent Variables	WJ Composite 12 months			WJ Composite 6 months			WJ Composite 3 months		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	405.85	.80	<.001	405.77	1.04	<.001	406.40	1.435	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.03	.02	.079
Region: West	-1.34	.87	.127	-1.70	1.06	.111	-2.56	1.36	.061
Region: Central East	.19	.91	.834	.43	1.11	.700	-.92	1.45	.527
Region: East	-.06	.93	.947	.60	1.11	.591	.24	1.45	.866
Male	-2.33	.39	<.001	-1.68	.58	.004	-2.25	.82	.007
Black	-3.04	.59	<.001	-2.27	.80	.005	-2.08	1.07	.052
Hispanic	-11.59	1.00	<.001	-13.46	1.54	<.001	-14.24	2.32	<.001
Native language other than English	-6.69	1.11	<.001	-6.66	1.63	<.001	-9.02	2.44	<.001
Time between start of school and testing	.13	.02	<.001	.18	.03	<.001	.18	.04	<.001
Has an IEP	-10.56	.69	<.001	-11.06	.99	<.001	-11.71	1.42	<.001
TN-VPK participation	**10.93**	.79	**<.001**	**10.74**	1.15	**<.001**	**10.84**	1.66	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. The composite is an average of the W-scores for the six subtests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table B-2
Statistical Models and Results for the WJ Letter-Word Subtest

Effect	WJ Letter-Word 12 months			WJ Letter-Word 6 months			WJ Letter-Word 3 months		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	328.77	1.41	<.001	329.04	1.78	<.001	331.26	2.47	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.05	.03	.052
Region: West	-.47	1.56	.762	-1.62	1.85	.381	-2.73	2.42	.263
Region: Central East	-.66	1.63	.686	-1.19	1.93	.538	-2.68	2.58	.299
Region: East	-1.71	1.66	.304	-.50	1.94	.796	-1.30	2.58	.613
Male	-4.02	.68	<.001	-4.32	.96	<.001	-6.14	1.38	<.001
Black	1.48	1.02	.148	3.84	1.36	.005	3.41	1.84	.064
Hispanic	-11.26	1.72	<.001	-13.04	2.58	<.001	-10.68	3.91	.006
Native language other than English	-2.71	1.91	.156	-4.59	2.73	.093	-6.38	4.10	.121
Time between start of school and testing	.24	.04	<.001	.32	.05	<.001	.38	.07	<.001
Has an IEP	-10.23	1.18	<.001	-11.47	1.65	<.001	-11.74	2.39	<.001
TN-VPK Participation	**22.02**	1.36	**<.001**	**21.83**	1.92	**<.001**	**19.85**	2.79	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table B-3
Statistical Models and Results for the WJ Spelling Subtest

Effect	WJ Spelling 12 months			WJ Spelling 6 months			WJ Spelling 3 months		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	362.69	1.20	<.001	362.87	1.67	<.001	362.83	2.35	<.001
Days from age eligibility cutoff	.06	<.01	<.001	.06	.01	<.001	.04	.03	.133
Region: West	1.57	1.27	.219	.61	1.69	.717	-1.41	2.24	.530
Region: Central East	.62	1.33	.640	1.69	1.76	.340	-.57	2.38	.810
Region: East	2.33	1.36	.088	4.15	1.78	.021	3.22	2.38	.178
Male	-5.98	.64	<.001	-4.93	.93	<.001	-5.35	1.36	<.001
Black	-1.57	.91	.086	-.01	1.29	.993	1.87	1.77	.291
Hispanic	-4.14	1.60	.010	-5.18	2.49	.038	-6.21	3.83	.106
Native language other than English	.43	1.77	.809	.95	2.63	.717	-.58	4.02	.884
Time between start of school and testing	.22	.04	<.001	.27	.05	<.001	.31	.07	<.001
Has an IEP	-11.32	1.10	<.001	-12.37	1.60	<.001	-12.79	2.35	<.001
TN-VPK Participation	**20.35**	1.28	**<.001**	**19.28**	1.86	**<.001**	**19.76**	2.75	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table B-4
Statistical Models and Results for the WJ Oral Comprehension Subtest

Effect	WJ Oral Comprehension 12 months			WJ Oral Comprehension 6 months			WJ Oral Comprehension 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	454.54	0.83	<.001	454.08	1.08	<.001	454.27	1.40	<.001
Days from age eligibility cutoff	.03	<.01	<.001	.03	.01	<.001	.01	.01	.414
Region: West	-3.39	.93	<.001	-3.01	1.16	.010	-3.10	1.36	.024
Region: Central East	1.32	.97	.175	1.79	1.21	.140	.61	1.44	.673
Region: East	-.87	.99	.377	-.88	1.22	.472	-.66	1.45	.648
Male	-1.53	.39	<.001	-.57	.56	.305	-.44	.79	.572
Black	-5.45	.60	<.001	-5.22	.81	<.001	-5.71	1.04	<.001
Hispanic	-11.09	1.00	<.001	-13.86	1.50	<.001	-15.39	2.23	<.001
Native language other than English	-12.35	1.11	<.001	-11.33	1.59	<.001	-12.34	2.34	<.001
Time between start of school and testing	.05	.02	.028	.09	.03	.005	.07	.04	.102
Has an IEP	-10.02	.68	<.001	-9.58	.96	<.001	-10.67	1.36	<.001
TN-VPK Participation	**3.25**	.79	**<.001**	**2.91**	1.11	**.009**	**4.03**	1.59	**.012**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table B-5
Statistical Models and Results for the WJ Picture Vocabulary Subtest

Effect	WJ Picture Vocabulary 12 months			WJ Picture Vocabulary 6 months			WJ Picture Vocabulary 3 months		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	465.47	.81	<.001	465.42	1.06	<.001	466.94	1.44	<.001
Days from age eligibility cutoff	.02	<.01	<.001	.02	.01	<.001	.02	.02	.159
Region: West	-1.87	.89	.038	-1.59	1.10	.150	-2.02	1.38	.146
Region: Central East	-.31	.93	.741	-.31	1.15	.785	-1.78	1.47	.229
Region: East	-.55	.95	.562	-.55	1.16	.638	-.98	1.47	.506
Male	.44	.40	.268	1.33	.57	.020	1.32	.82	.109
Black	-1.78	.59	.003	-2.28	.81	.005	-3.03	1.08	.005
Hispanic	-18.54	1.00	<.001	-20.87	1.54	<.001	-20.85	2.33	<.001
Native language other than English	-13.84	1.11	<.001	-13.57	1.63	<.001	-18.25	2.44	<.001
Time between start of school and testing	.03	.02	.249	.04	.03	.182	.03	.04	.548
Has an IEP	-7.92	.69	<.001	-7.60	.99	<.001	-7.57	1.43	<.001
TN-VPK Participation	**3.69**	.80	**<.001**	**3.69**	1.14	**.001**	2.69	1.67	.107

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table B-6
Statistical Models and Results for the WJ Applied Problems Subtest

Effect	WJ Applied Problems 12 months			WJ Applied Problems 6 months			WJ Applied Problems 3 months		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	407.91	1.26	<.001	407.52	1.55	<.001	407.86	2.16	<.001
Days from age eligibility cutoff	.05	<.01	<.001	.04	.01	<.001	.02	.02	.451
Region: West	-3.03	1.36	.027	-3.17	1.52	.039	-3.50	2.06	.092
Region: Central East	.10	1.42	.943	.54	1.59	.735	-.22	2.20	.920
Region: East	-.01	1.45	.995	-.28	1.60	.862	-.81	2.20	.712
Male	-2.02	.64	.002	-.56	.90	.532	-1.97	1.24	.114
Black	-8.07	.94	<.001	-8.16	1.21	<.001	-7.98	1.62	<.001
Hispanic	-17.66	1.61	<.001	-19.09	2.40	<.001	-24.34	3.51	<.001
Native language other than English	-9.90	1.79	<.001	-9.95	2.53	<.001	-12.58	3.68	<.001
Time between start of school and testing	.14	.04	<.001	.20	.05	<.001	.14	.06	.023
Has an IEP	-15.93	1.11	<.001	-16.87	1.54	<.001	-18.60	2.15	<.001
TN-VPK Participation	**8.37**	1.28	**<.001**	**9.80**	1.79	**<.001**	**11.36**	2.51	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table B-7
Statistical Models and Results for the WJ Quantitative Concepts Subtest

Effect	WJ Quantitative Concepts 12 months			WJ Quantitative Concepts 6 months			WJ Quantitative Concepts 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	415.58	0.78	<.001	415.42	1.03	<.001	415.21	1.39	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.02	.02	.132
Region: West	-.77	.84	.360	-1.43	1.03	.169	-2.75	1.28	.034
Region: Central East	.22	.88	.801	.20	1.08	.853	-.82	1.37	.551
Region: East	.63	.90	.483	1.72	1.08	.114	1.97	1.37	.153
Male	-.91	.40	.021	-1.00	.58	.087	-.85	.83	.309
Black	-2.66	.58	<.001	-1.48	.80	.064	-.93	1.05	.380
Hispanic	-6.86	1.00	<.001	-8.78	1.56	<.001	-8.12	2.33	<.001
Native language other than English	-1.81	1.11	.103	-1.29	1.65	.433	-3.96	2.45	.106
Time between start of school and testing	.10	.02	<.001	.14	.03	<.001	.16	.04	<.001
Has an IEP	-7.94	.69	<.001	-8.55	1.00	<.001	-9.16	1.44	<.001
TN-VPK Participation	**7.88**	.80	**<.001**	**7.08**	1.16	**<.001**	**7.45**	1.68	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table C-1
Robustness Checks for Models Estimating the TN-VPK Effect

Outcome and Sample	12 months			6 months			3 months		
	(1) SD[1]	(2) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ2)/(1)[2]	(3) SD[1]	(4) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ4)/(3)[2]	(5) SD[1]	(6) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ6)/(5)[2]
WJ Composite									
All cases	19.34	10.93 (.79)		16.97	10.74 (1.15)		16.50	10.84 (1.66)	
Outermost 1% trimmed		10.97 (.79)	-.002		10.30 (1.16)	.026		11.36 (1.67)	-.032
Outermost 5% trimmed		10.93 (.81)	.000		9.70 (1.20)	.061		10.50 (1.77)	.021
Outermost 10% trimmed		10.59 (.83)	.018		9.18 (1.28)	.093		9.17 (1.85)	.101
WJ Letter-Word									
All cases	29.71	22.02 (1.36)		27.24	21.83(1.92)		26.95	19.85 (2.79)	
Outermost 1% trimmed		22.11 (1.37)	-.003		21.17 (1.94)	.024		20.56 (2.80)	-.026
Outermost 5% trimmed		22.25 (1.40)	.008		20.34 (2.02)	.055		19.26 (2.96)	.022
Outermost 10% trimmed		21.82 (1.43)	.007		18.62 (2.16)	.118		16.33 (3.10)	.131
WJ Spelling									
All cases	30.00	20.34 (1.28)		26.61	19.28 (1.86)		25.73	19.76 (2.75)	
Outermost 1% trimmed		20.52 (1.28)	.011		18.58 (1.87)	.026		20.76 (2.75)	-.039
Outermost 5% trimmed		20.20 (1.31)	.021		17.99 (1.95)	.048		19.32 (2.89)	.017
Outermost 10% trimmed		19.99 (1.34)	.028		17.63 (2.08)	.062		18.91 (3.03)	.033

Appendix Table C-1 (*continued*)
Robustness Checks for Models Estimating the TN-VPK Effect

Outcome and Sample	12 months			6 months			3 months		
	(1) SD[1]	(2) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ2)/(1)[2]	(3) SD[1]	(4) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ4)/(3)[2]	(5) SD[1]	(6) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ6)/(5)[2]
WJ Oral Comprehension									
All cases	16.45	3.25 (.79)		15.19	2.91 (1.11)		14.85	4.03 (1.59)	
Outermost 1% trimmed		3.17 (.79)	.005		2.90 (1.13)	<.001		4.48 (1.61)	-.030
Outermost 5% trimmed		3.15 (.81)	.006		2.47 (1.17)	.029		3.93 (1.69)	.007
Outermost 10% trimmed		2.95 (.83)	.018		2.20 (1.24)	.047		3.09 (1.79)	.063
WJ Picture Vocabulary									
All cases	16.34	3.69 (.80)		15.92	3.69 (1.14)		16.33	2.69 (1.67)	
Outermost 1% trimmed		3.80 (.80)	-.007		3.36 (1.16)	.021		3.03 (1.67)	-.039
Outermost 5% trimmed		3.89 (.82)	-.012		2.46 (1.22)	.077		1.91 (1.80)	.048
Outermost 10% trimmed		3.75 (.84)	-.004		2.05 (1.31)	.103		1.06 (1.90)	.100
WJ Applied Problems	27.18			23.87			23.56		
All cases		8.37 (1.28)			9.80 (1.79)			11.36 (2.51)	
Outermost 1% trimmed		8.36 (1.28)	<.001		9.24 (1.80)	.023		11.61 (2.52)	-.011
Outermost 5% trimmed		8.23 (1.32)	.006		8.75 (1.89)	.044		11.19 (2.66)	.007
Outermost 10% trimmed		7.80 (1.34)	.021		8.83 (2.00)	.041		9.29 (2.83)	.088

Appendix Table C-1 (*continued*)
Robustness Checks for Models Estimating the TN-VPK Effect

Outcome and Sample	12 months			6 months			3 months		
	(1) SD[1]	(2) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ2)/(1)[2]	(3) SD[1]	(4) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ4)/(3)[2]	(5) SD[1]	(6) Coefficient for TN-VPK vs. No Pre-K (SE)	Difference in Estimates (Δ6)/(5)[2]
WJ Quantitative Concepts									
All cases	17.32	7.88 (.80)		15.41	7.08 (1.16)		14.89	7.45 (1.68)	
Outermost 1% trimmed		7.84 (.80)	.002		6.67 (1.17)	.027		7.70 (1.70)	-.017
Outermost 5% trimmed		7,87 (.82)	<.001		6.29 (1.22)	.051		7.30 (1.77)	.010
Outermost 10% trimmed		7.23 (.84)	.038		5.84 (1.30)	.080		6.34 (1.86)	.075
Average effect size for:[3]									
All cases vs. 1% trimmed		.004			.021			.028	
All cases vs. 5% trimmed		.008			.052			.019	
All cases vs. 10% trimmed		.019			.078			.084	

Note. Only children eligible for free or reduced price lunch were included.

[1] This is the standard deviation for the TN-VPK sample, i.e., children who had completed TN-VPK and were tested at the beginning of kindergarten.

[2] Effect sizes for the difference between the coefficients of the sample with all cases and the respective trimmed sample. The coefficient differences are computed from columns (2), (4), and (6) respectively by subtracting each trimmed value from the all cases value in turn. For comparison purposes, those differences (Δ) are standardized into an effect size metric by dividing each by the standard deviation for all cases shown in columns (1), (3), and (5) respectively. Those standardized differences are then reported in the "Difference in Estimates" columns.

[3] This is the average of the absolute values of the "Difference in Estimates" effect sizes for the respective comparison of the two samples.

Appendix Table D-1
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Male and Female Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Composite								
12-month bandwidth					12.92			.464
Female	412.90	1044	402.25	1044		10.65	.824	
Male	410.85	1034	399.63	1022		11.22	.868	
6-month bandwidth					12.66			.984
Female	412.78	511	402.05	553		10.73	.848	
Male	411.11	498	400.36	505		10.75	.849	
3 -month bandwidth					11.94			.593
Female	412.66	267	402.25	181		10.41	.872	
Male	410.87	252	399.58	259		11.29	.946	
Letter-Word								
12-month bandwidth					21.15			
Female	348.84	1044	326.60	1044		22.24	1.052	.731
Male	344.58	1034	322.81	1022		21.77	1.029	
6-month bandwidth					19.71			.144
Female	349.65	511	326.44	553		23.21	1.178	
Male	343.88	498	323.46	505		20.42	1.036	
3 -month bandwidth					19.17			.645
Female	348.74	267	328.27	181		20.47	1.068	
Male	341.93	252	322.74	259		19.19	1.001	

Appendix Table D-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Male and Female Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Spelling								
12-month bandwidth					20.82			.335
Female	382.77	1044	361.84	1044		20.93	1.005	
Male	376.19	1034	356.47	1022		19.72	.947	
6-month bandwidth					20.62			.181
Female	382.74	511	362.23	553		20.51	.995	
Male	376.53	498	358.50	505		18.03	.874	
3 -month bandwidth					20.43			.322
Female	382.56	267	361.49	272		21.07	1.031	
Male	375.80	252	357.43	259		18.37	.899	
Oral Comprehension								
12-month bandwidth					13.71			.079
Female	451.96	1044	449.37	1044		2.59	.189	
Male	451.12	1034	447.16	1022		3.96	.289	
6-month bandwidth					13.08			.004
Female	450.53	511	449.54	553		.99	.076	
Male	451.98	498	447.10	505		4.88	.373	
3 -month bandwidth					12.87			.052
Female	451.57	267	449.03	272		2.54	.197	
Male	452.73	252	447.13	259		5.60	.435	

Appendix Table D-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Male and Female Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Picture Vocabulary								
12-month bandwidth					10.72			.572
Female	464.42	1044	460.94	1044		3.48	.325	
Male	465.08	1034	461.16	1022		3.92	.366	
6-month bandwidth					10.89			.459
Female	464.00	511	460.73	553		3.27	.300	
Male	465.77	498	461.66	505		4.11	.377	
3 -month bandwidth					10.41			.568
Female	463.60	267	461.37	272		2.23	.214	
Male	465.42	252	462.24	259		3.18	.305	
Applied Problems								
12-month bandwidth					17.74			.096
Female	408.51	1044	401.16	1044		7.35	.414	
Male	407.55	1034	398.08	1022		9.47	.534	
6-month bandwidth					18.12			.767
Female	409.48	511	399.94	553		9.54	.526	
Male	409.19	498	399.13	505		10.06	.555	
3 -month bandwidth					16.08			.075
Female	409.58	267	400.39	272		9.19	.572	
Male	409.94	252	396.31	259		13.63	.848	

Appendix Table D-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Male and Female Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[d]	*p*[d]
	Mean[a]	*N*	Mean[a]	*N*				
Quantitative Concepts								
12-month bandwidth					13.99			.146
Female	420.69	1044	413.56	1044		7.13	.510	
Male	420.55	1034	412.08	1022		8.47	.605	
6-month bandwidth					13.43			.807
Female	420.26	511	413.33	553		6.93	.516	
Male	419.41	498	412.20	505		7.21	.537	
3 -month bandwidth					12.49			.596
Female	419.89	267	412.87	272		7.02	.562	
Male	419.51	252	411.60	259		7.91	.633	

Note. Only children eligible for free or reduced price lunch were included. Previous analyses showed that the appropriate functional form was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Tables E-1 to E-7 report the full results for each outcome and bandwidth.

[a] Reported are the marginal means from the multilevel model.

[b] To facilitate comparison of these effect sizes with those reported for the overall pre-k effect, effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the standard deviation of the treatment (TN-VPK) group for the sample children eligible for free or reduced price lunch.

[c] This is the *p*-value for the interaction term for condition (TN-VPK vs. No Pre-K) and whether the child was Male or Female. Estimates were based on a multilevel model with children nested within pre-k classroom. In addition to condition, gender, and days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) whether the child was Black; (3) whether the child was Hispanic; (4) whether the child's native language was not English; (5) whether the child had an IEP placement; (6) the number of days elapsed between the Woodcock-Johnson testing date and the start of school (centered at the grand mean); and (7) the interaction term.

Appendix Table E-1

Statistical Models for the Effects of TN-VPK on the WJ Composite: Comparison of Male and Female Children

Effect	WJ Composite 12 months			WJ Composite 6 months			WJ Composite 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	406.00	.82	<.001	405.78	1.08	<.001	406.62	1.48	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.03	.02	.078
Region: West	-1.33	.87	.128	-1.70	1.06	.111	-2.55	1.36	.062
Region: Central East	.18	.91	.841	.43	1.11	.700	-.92	1.45	.525
Region: East	-.07	.93	.946	.60	1.11	.591	.25	1.45	.862
Male	-2.62	.55	<.001	-1.69	.80	.034	-2.67	1.14	.019
Black	-3.05	.59	<.001	-2.27	.80	.005	-2.13	1.08	.048
Hispanic	-11.59	1.00	<.001	-13.46	1.54	<.001	-14.25	2.32	<.001
Native language other than English	-6.69	1.11	<.001	-6.66	1.63	<.001	-9.00	2.44	<.001
Time between start of school and testing	.13	.02	<.001	.18	.03	<.001	.18	.04	<.001
Has an IEP	-10.55	.69	<.001	-11.06	.99	<.001	11.74	1.42	<.001
TN-VPK participation	10.65	.88	<.001	10.73	1.28	<.001	10.41	1.85	<.001
TN-VPK participation x Male	.57	.78	.464	.02	1.14	.984	.88	1.65	.593

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. The composite is an average of the W-scores for the six subtests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table E-2
Statistical Models for the Effects of TN-VPK on the WJ Letter-Word Subtest: Comparison of Male and Female Children

Effect	WJ Letter-Word 12 months			WJ Letter-Word 6 months			WJ Letter-Word 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	328.66	1.45	<.001	328.36	1.84	<.001	330.95	2.57	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.05	.03	.053
Region: West	-.48	1.56	.762	-1.62	1.85	.380	-2.74	2.42	.259
Region: Central East	-.66	1.63	.689	-1.15	1.93	.553	-2.68	2.57	.300
Region: East	-1.71	1.66	.305	-.49	1.94	.802	-1.31	2.58	.611
Male	-3.79	.95	<.001	-2.98	1.33	.025	-5.53	1.91	.004
Black	1.48	1.02	.147	3.87	1.36	.005	3.48	1.84	.060
Hispanic	-11.26	1.72	<.001	-13.04	2.58	<.001	-10.65	3.91	.007
Native language other than English	-2.70	1.91	.157	-4.51	2.73	.099	-6.42	4.10	.118
Time between start of school and testing	.24	.04	<.001	.32	.05	<.001	.37	.07	<.001
Has an IEP	-10.24	1.18	<.001	-11.47	1.65	<.001	-11.70	2.39	<.001
TN-VPK participation	22.24	1.51	<.001	23.22	2.14	<.001	20.46	3.10	<.001
TN-VPK participation x Male	-.46	1.35	.731	-2.79	1.91	.144	-1.27	2.77	.645

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table E-3
Statistical Models for the Effects of TN-VPK on the WJ Spelling Subtest: Comparison of Male and Female Children

Effect	WJ Spelling 12 months			WJ Spelling 6 months			WJ Spelling 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	362.40	1.24	<.001	362.27	1.73	<.001	362.15	2.44	<.001
Days from age eligibility cutoff	.06	<.01	<.001	.06	.01	<.001	.04	.03	.134
Region: West	1.57	1.27	.219	.61	1.69	.717	-1.42	2.23	.522
Region: Central East	.64	1.33	.630	1.73	1.77	.330	-.56	2.38	.815
Region: East	2.34	1.36	.086	4.16	1.78	.021	3.20	2.38	.181
Male	-5.97	.89	<.001	-3.74	1.29	.004	-4.06	1.88	.031
Black	-1.56	.91	.089	.01	1.29	.997	2.01	1.77	.256
Hispanic	-4.14	1.60	.010	-5.19	2.49	.037	-6.16	3.83	.108
Native language other than English	.44	1.77	.804	1.03	2.63	.698	-.66	4.02	.870
Time between start of school and testing	.22	.04	<.001	.27	.05	<.001	.31	.07	<.001
Has an IEP	-11.35	1.10	<.001	-12.37	1.60	<.001	-12.71	2.35	<.001
TN-VPK participation	20.93	1.41	<.001	20.50	2.07	<.001	21.07	3.05	<.001
TN-VPK participation x Male	-1.22	1.26	.335	-2.47	1.85	.181	-2.70	2.72	.322

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table E-4
Statistical Models for the Effects of TN-VPK on the WJ Oral Comprehension Subtest: Comparison of Male and Female Children

Effect	WJ Oral Comprehension 12 months			WJ Oral Comprehension 6 months			WJ Oral Comprehension 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	454.68	0.85	<.001	455.01	1.12	<.001	455.03	1.45	<.001
Days from age eligibility cutoff	.03	<.01	<.001	.03	.01	<.001	.01	.01	.409
Region: West	-3.39	.93	<.001	-3.01	1.16	.010	-3.06	1.35	.026
Region: Central East	1.31	.97	.181	1.74	1.21	.154	.60	1.44	.680
Region: East	-.88	.99	.373	-.89	1.22	.469	-.64	1.45	.660
Male	-2.21	.55	<.001	-2.44	.77	.002	-1.90	1.09	.080
Black	-5.47	.60	<.001	-5.24	.81	<.001	-5.88	1.04	<.001
Hispanic	-11.09	.99	<.001	-13.85	1.50	<.001	-15.45	2.23	<.001
Native language other than English	-12.36	1.11	<.001	-11.44	1.59	<.001	-12.25	2.33	<.001
Time between start of school and testing	.05	.02	.030	.09	.03	.005	.07	.04	.090
Has an IEP	-10.00	.68	<.001	-9.59	.96	<.001	-10.75	1.36	<.001
TN-VPK participation	2.59	.88	<.001	2.99	1.24	.424	2.54	1.77	.151
TN-VPK participation x Male	1.37	.78	.079	**3.88**	1.10	**<.001**	**3.06**	1.58	**.052**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table E-5
Statistical Models for the Effects of TN-VPK on the WJ Picture Vocabulary Subtest: Comparison of Male and Female Children

Effect	WJ Picture Vocabulary 12 months			WJ Picture Vocabulary 6 months			WJ Picture Vocabulary 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	465.58	.83	<.001	465.63	1.10	<.001	467.17	1.50	<.001
Days from age eligibility cutoff	.02	<.01	<.001	.02	.01	<.001	.02	.02	.158
Region: West	-1.87	.89	.038	-1.59	1.10	.150	-2.01	1.38	.158
Region: Central East	-.32	.93	.736	-.33	1.15	.777	-1.79	1.47	.228
Region: East	-.55	.95	.561	-.55	1.16	.637	-.98	1.48	.510
Male	.22	.56	.698	.93	.79	.242	.87	1.14	.444
Black	-1.78	.59	.003	-2.28	.81	.005	-3.09	1.08	.005
Hispanic	-18.54	1.00	<.001	-20.87	1.54	<.001	-20.87	2.33	<.001
Native language other than English	-13.84	1.11	<.001	-13.59	1.63	<.001	-18.23	2.44	<.001
Time between start of school and testing	.03	.02	.252	.04	.03	.184	.03	.04	.537
Has an IEP	-7.91	.69	<.001	-7.60	.99	<.001	-7.60	1.43	<.001
TN-VPK participation	3.48	.88	<.001	3.27	1.28	.011	2.23	1.85	.230
TN-VPK participation x Male	.45	.79	.572	.84	1.14	.459	.95	1.65	.568

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table E-6
Statistical Models for the Effects of TN-VPK on the WJ Applied Problems Subtest: Comparison of Male and Female Children

Effect	WJ Applied Problems 12 months			WJ Applied Problems 6 months			WJ Applied Problems 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	408.42	1.29	<.001	407.65	1.61	<.001	408.96	2.24	<.001
Days from age eligibility cutoff	.05	<.01	<.001	.04	.01	<.001	.02	.02	.447
Region: West	-3.03	1.36	.027	-3.17	1.52	.039	-3.44	2.05	.096
Region: Central East	.07	1.42	.959	.53	1.59	.739	-.24	2.19	.912
Region: East	-.02	1.45	.991	-.28	1.60	.861	-.78	2.19	.724
Male	-3.07	.90	<.001	-.82	1.24	.511	-4.07	1.72	.018
Black	-8.10	.94	<.001	-8.16	1.21	<.001	-8.24	1.62	<.001
Hispanic	-17.65	1.61	<.001	-19.09	2.40	<.001	-24.44	3.51	<.001
Native language other than English	-9.92	1.79	<.001	-9.97	2.53	<.001	-12.44	3.68	<.001
Time between start of school and testing	.14	.04	<.001	.20	.05	<.001	.15	.06	.020
Has an IEP	-15.90	1.11	<.001	-16.87	1.54	<.001	-18.71	2.15	<.001
TN-VPK participation	7.36	1.42	<.001	9.54	2.00	<.001	9.19	2.79	.001
TN-VPK participation x Male	2.11	1.27	.096	**.53**	1.79	.767	4.43	2.49	.075

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table E-7
Statistical Models for the Effects of TN-VPK on the WJ Quantitative Concepts Subtest: Comparison of Male and Female Children

Effect	WJ Quantitative Concepts 12 months			WJ Quantitative Concepts 6 months			WJ Quantitative Concepts 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	415.86	0.80	<.001	415.49	1.07	<.001	415.43	1.45	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.02	.02	.131
Region: West	-.77	.84	.362	-1.43	1.03	.169	-2.74	1.29	.035
Region: Central East	.21	.88	.815	.20	1.08	.856	-.83	1.38	.548
Region: East	.63	.90	.486	1.72	1.09	.115	1.98	1.37	.152
Male	-.1.48	.56	.008	-1.13	.81	.160	-1.27	1.15	.270
Black	-2.67	.58	<.001	-1.48	.80	.064	-.97	1.06	.358
Hispanic	-6.85	1.00	<.001	-8.78	1.56	<.001	-8.13	2.34	<.001
Native language other than English	-1.82	1.11	.101	-1.30	1.65	.430	-3.95	2.45	.107
Time between start of school and testing	.10	.02	<.001	.14	.03	<.001	.16	.04	<.001
Has an IEP	-7.92	.69	<.001	-8.55	1.00	<.001	-9.19	1.44	<.001
TN-VPK participation	7.33	.88	<.001	6.94	1.30	<.001	7.02	1.87	<.001
TN-VPK participation x Male	1.15	.79	.146	.28	1.16	.807	.88	1.67	.596

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table F-1
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Black and Non-Black Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Composite								
12-month bandwidth					12.92			.813
Black	410.01	750	398.94	748		11.07	.857	
Not Black	412.95	1328	402.08	1318		10.87	.841	
6-month bandwidth					12.66			.931
Black	410.56	372	399.74	376		10.82	.855	
Not Black	412.78	737	402.06	682		10.72	.847	
3 -month bandwidth					11.94			.779
Black	410.28	197	399.81	181		10.47	.877	
Not Black	412.62	322	401.63	350		10.99	.920	
Letter-Word								
12-month bandwidth					21.15			
Black	347.20	750	326.11	748		21.09	.977	.371
Not Black	346.39	1328	323.99	1318		22.40	1.059	
6-month bandwidth					19.71			.228
Black	348.37	372	328.33	376		20.04	1.017	
Not Black	345.83	637	323.26	682		22.57	1.145	
3 -month bandwidth					19.17			.486
Black	346.88	197	328.56	181		18.32	.956	
Not Black	344.52	322	324.08	350		20.44	1.066	

Appendix Table F-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Black and Non-Black Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Spelling								
12-month bandwidth					20.82			.156
Black	379.22	750	357.49	748		21.73	1.043	
Not Black	379.79	1328	360.01	1318		19.78	.950	
6-month bandwidth					20.62			.518
Black	380.16	372	359.96	376		20.20	.980	
Not Black	379.50	637	360.60	682		18.90	.917	
3 -month bandwidth					20.43			.920
Black	380.36	197	360.82	181		19.54	.956	
Not Black	378.64	322	358.80	350		19.84	.971	
Oral Comprehension								
12-month bandwidth					13.71			.606
Black	448.21	750	444.65	748		3.56	.260	
Not Black	453.44	1328	450.32	1318		3.12	.228	
6-month bandwidth					13.08			.632
Black	448.15	372	444.82	376		3.33	.255	
Not Black	453.07	637	450.33	682		2.74	.209	
3 -month bandwidth					12.87			.701
Black	448.71	197	444.20	181		4.51	.350	
Not Black	454.09	322	450.25	350		3.84	.298	

Appendix Table F-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Black and Non-Black Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Picture Vocabulary								
12-month bandwidth					10.72			.036
Black	462.95	750	460.54	748		2.41	.225	
Not Black	465.66	1328	461.45	1318		4.21	.393	
6-month bandwidth					10.89			.052
Black	462.55	372	460.58	376		1.97	.181	
Not Black	466.07	637	461.67	682		4.40	.404	
3 -month bandwidth					10.41			.025
Black	461.16	197	461.39	181		-.23	-.022	
Not Black	466.20	322	462.37	350		3.83	.368	
Applied Problems								
12-month bandwidth					17.74			.498
Black	403.21	750	394.17	748		9.04	.510	
Not Black	410.80	1328	402.70	1318		8.10	.457	
6-month bandwidth					18.12			.405
Black	404.72	372	393.77	376		10.95	.604	
Not Black	412.04	637	402.72	682		9.32	.514	
3 -month bandwidth					16.08			.870
Black	404.81	197	393.13	181		11.68	.726	
Not Black	412.57	322	401.34	350		11.23	.698	

Appendix Table F-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Black and Non-Black Children

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[d]	*p*[d]
	Mean[a]	*N*	Mean[a]	*N*				
Quantitative Concepts								
12-month bandwidth					13.99			.261
Black	419.37	750	410.80	748		8.57	.613	
Not Black	421.53	1328	413.93	1318		7.60	.543	
6-month bandwidth					13.43			.101
Black	419.65	372	411.11	376		8.54	.636	
Not Black	420.06	637	413.60	682		6.46	.481	
3 -month bandwidth					12.49			.201
Black	419.91	197	410.78	181		9.13	.731	
Not Black	419.67	322	412.89	350		6.78	.543	

Note. Only children eligible for free or reduced price lunch were included. Previous analyses showed that the appropriate functional form was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Appendix Tables G-1-to G-7 report the full results for each outcome and bandwidth.

[a] Reported are the marginal means from the multilevel model.

[b] To facilitate comparison of these effect sizes with those reported for the overall pre-k effect, effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the standard deviation of the treatment (TN-VPK) group for the sample of children eligible for free or reduced price lunch.

[c] This is the *p*-value for the interaction term for condition (TN-VPK vs. No Pre-K) and whether the child was Black or not Black. Estimates were based on a multilevel model with children nested within pre-k classroom. In addition to condition, Black, and days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) whether the child was Male; (3) whether the child was Hispanic; (4) whether the child's native language was not English; (5) whether the child had an IEP placement; (6) the number of days elapsed between the Woodcock-Johnson testing date and the start of school (centered at the grand mean); and (7) the interaction term.

Appendix Table G-1

Statistical Models for the Effects of TN-VPK on the WJ Composite: Comparison of Black and Non-Black Children

Effect	WJ Composite 12 months			WJ Composite 6 months			WJ Composite 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	405.88	.81	<.001	405.79	1.05	<.001	406.31	1.46	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.03	.02	.078
Region: West	-1.34	.87	.128	-1.70	1.06	.111	-2.56	1.36	.062
Region: Central East	.19	.91	.833	.43	1.11	.700	-.91	1.49	.532
Region: East	-.06	.93	.948	.60	1.11	.591	.26	1.45	.857
Male	-2.33	.39	<.001	-1.68	.58	.004	-2.23	.83	.007
Black	-3.14	.72	<.001	-2.32	1.01	.021	-1.83	1.41	.196
Hispanic	-11.59	1.00	<.001	-13.45	1.54	<.001	-14.25	2.32	<.001
Native language other than English	-6.69	1.11	<.001	-6.66	1.63	<.001	-8.98	2.44	<.001
Time between start of school and testing	.13	.02	<.001	.18	.03	<.001	.18	.04	<.001
Has an IEP	-10.56	.69	<.001	-11.06	.99	<.001	11.71	1.42	<.001
TN-VPK participation	10.87	.83	<.001	10.71	1.21	<.001	10.99	1.74	<.001
TN-VPK participation x Black	.20	.85	.813	.11	1.25	.931	-.51	1.81	.779

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. The composite is an average of the W-scores for the six subtests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table G-2
Statistical Models for the Effects of TN-VPK on the WJ Letter-Word Subtest: Comparison of Black and Non-Black Children

Effect	WJ Letter-Word 12 months			WJ Letter-Word 6 months			WJ Letter-Word 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	328.59	1.42	<.001	328.69	1.80	<.001	330.91	2.52	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.05	.03	.051
Region: West	-.48	1.56	.761	-1.64	1.85	.377	-2.69	2.42	.268
Region: Central East	-.67	1.63	.683	-1.18	1.93	.541	-2.64	2.57	.307
Region: East	-1.72	1.66	.303	-.50	1.94	.797	-1.23	2.58	.063
Male	-4.02	.68	<.001	-4.33	.96	<.001	-6.06	1.38	<.001
Black	2.12	1.25	.089	5.07	1.70	.003	4.48	2.40	.062
Hispanic	-11.28	1.72	<.001	-13.12	2.58	<.001	-10.73	3.91	.006
Native language other than English	-2.67	1.91	.161	-4.43	2.74	.106	-6.21	4.11	.131
Time between start of school and testing	.24	.04	<.001	.32	.05	<.001	.38	.07	<.001
Has an IEP	-10.23	1.18	<.001	-11.47	1.65	<.001	-11.73	2.39	<.001
TN-VPK participation	22.40	1.43	<.001	22.57	2.02	<.001	20.44	2.92	<.001
TN-VPK participation x Black	-1.31	1.47	.371	-2.52	2.09	.228	-2.12	33.04	.486

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table G-3
Statistical Models for the Effects of TN-VPK on the WJ Spelling Subtest: Comparison of Black and Non-Black Children

Effect	WJ Spelling 12 months			WJ Spelling 6 months			WJ Spelling 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	362.97	1.22	<.001	363.05	1.69	<.001	362.78	2.40	<.001
Days from age eligibility cutoff	.06	<.01	<.001	.06	.01	<.001	.04	.03	.133
Region: West	1.57	1.27	.220	.62	1.69	.715	-1.40	2.24	.532
Region: Central East	.63	1.33	.636	1.68	1.76	.341	-.57	2.38	.813
Region: East	2.34	1.36	.087	4.15	1.78	.021	3.23	2.39	.177
Male	-5.99	.64	<.001	-4.92	.93	<.001	-5.33	1.36	<.001
Black	-2.52	1.13	.026	-.65	1.62	.689	2.02	2.33	.386
Hispanic	-4.11	1.60	.010	-5.14	2.49	.039	-6.22	3.84	.105
Native language other than English	.38	1.77	.831	.87	2.64	.742	-.56	4.03	.889
Time between start of school and testing	.22	.04	<.001	.27	.05	<.001	.31	.07	<.001
Has an IEP	-11.33	1.11	<.001	-12.37	1.60	<.001	-12.79	2.35	<.001
TN-VPK participation	19.78	1.34	<.001	18.89	1.95	<.001	19.84	2.88	<.001
TN-VPK participation x Black	1.95	1.37	.156	1.31	2.02	.518	-.30	2.99	.920

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table G-4
Statistical Models for the Effects of TN-VPK on the WJ Oral Comprehension Subtest: Comparison of Black and Non-Black Children

Effect	WJ Oral Comprehension 12 months			WJ Oral Comprehension 6 months			WJ Oral Comprehension 3 months		
	b	SE	*p*	*b*	SE	*p*	*B*	SE	*p*
Intercept	454.61	0.84	<.001	454.16	1.10	<.001	454.38	1.43	<.001
Days from age eligibility cutoff	.03	<.01	<.001	.03	.01	<.001	.01	.01	.417
Region: West	-3.39	.93	<.001	-3.01	1.16	.010	-3.11	1.36	.023
Region: Central East	1.33	.97	.174	1.79	1.21	.140	.60	1.44	.681
Region: East	-.87	.99	.378	-.88	1.21	.306	-.68	1.45	.637
Male	-1.53	.39	<.001	-.57	.56	.306	-.47	.79	.553
Black	-5.67	.73	<.001	-5.51	1.00	<.001	-6.05	1.36	<.001
Hispanic	-11.09	1.00	<.001	-13.84	1.50	<.001	-15.37	2.23	<.001
Native language other than English	-12.36	1.11	<.001	-11.37	1.59	<.001	-12.39	2.34	<.001
Time between start of school and testing	.05	.02	.030	.09	.03	.006	.07	.04	.106
Has an IEP	-10.02	.68	<.001	-9.58	.96	<.001	-10.67	1.36	<.001
TN-VPK participation	3.12	.83	<.001	2.74	1.17	.019	3.84	1.67	.021
TN-VPK participation x Black	.44	.85	.606	.58	1.21	.632	.67	1.73	.701

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table G-5
Statistical Models for the Effects of TN-VPK on the WJ Picture Vocabulary Subtest: Comparison of Black and Non-Black Children

Effect	WJ Picture Vocabulary 12 months			WJ Picture Vocabulary 6 months			WJ Picture Vocabulary 3 months		
	b	SE	*p*	*b*	SE	*p*	*B*	SE	*p*
Intercept	465.22	.82	<.001	465.09	1.08	<.001	466.26	1.47	<.001
Days from age eligibility cutoff	.02	<.01	<.001	.02	.01	<.001	.02	.02	.148
Region: West	-1.87	.89	.037	-1.61	1.10	.147	-1.96	1.38	.158
Region: Central East	-.32	.93	.733	-.30	1.15	.791	-1.70	1.47	.250
Region: East	-.56	.95	.555	-.54	1.16	.642	-.85	1.47	.563
Male	.45	.40	.257	1.33	.57	.020	1.47	.83	.075
Black	-.90	.72	.213	-1.09	1.01	.279	-.98	1.42	.489
Hispanic	-18.57	1.00	<.001	-20.95	1.54	<.001	-20.96	2.33	<.001
Native language other than English	-13.79	1.11	<.001	-13.41	1.63	<.001	-17.93	2.44	<.001
Time between start of school and testing	.03	.02	.201	.05	.03	.142	.03	.04	.480
Has an IEP	-7.91	.69	<.001	-7.60	.98	<.001	-7.56	1.42	<.001
TN-VPK participation	4.21	.84	<.001	4.40	1.20	<.001	3.84	1.74	.028
TN-VPK participation x Black	**-1.80**	.86	**.036**	**-2.42**	1.25	**.052**	**-4.06**	1.81	**.025**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table G-6
Statistical Models for the Effects of TN-VPK on the WJ Applied Problems Subtest: Comparison of Black and Non-Black Children

Effect	WJ Applied Problems 12 months			WJ Applied Problems 6 months			WJ Applied Problems 3 months		
	b	SE	*p*	*b*	SE	*p*	*B*	SE	*p*
Intercept	408.04	1.27	<.001	407.75	1.57	<.001	407.93	2.21	<.001
Days from age eligibility cutoff	.05	<.01	<.001	.04	.01	<.001	.02	.02	.453
Region: West	-3.03	1.36	.027	-3.17	1.52	.039	-3.51	2.06	.091
Region: Central East	.11	1.42	.940	.53	1.59	.737	-.23	2.20	.917
Region: East	<.01	1.45	.998	-.28	1.59	.859	-.83	2.20	.707
Male	-2.03	.64	.002	-.56	.90	.533	-1.99	1.25	.112
Black	-8.53	1.15	<.001	-8.95	1.54	<.001	-8.21	2.13	<.001
Hispanic	-17.64	1.61	<.001	-19.04	2.40	<.001	-24.32	3.51	<.001
Native language other than English	-9.92	1.79	<.001	-10.06	2.53	<.001	-12.62	3.69	<.001
Time between start of school and testing	.13	.04	<.001	.20	.05	<.001	.15	.06	.024
Has an IEP	-15.94	1.11	<.001	-16.87	1.54	<.001	-18.60	2.15	<.001
TN-VPK participation	8.10	1.34	<.001	9.32	1.89	<.001	11.23	2.63	.001
TN-VPK participation x Black	.94	1.38	.498	1.62	1.95	.405	.45	2.74	.870

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table G-7
Statistical Models for the Effects of TN-VPK on the WJ Quantitative Problems Subtest: Comparison of Black and Non-Black Children

Effect	WJ Quantitative Concepts 12 months			WJ Quantitative Concepts 6 months			WJ Quantitative Concepts 3 months		
	b	SE	*p*	*b*	SE	*p*	*B*	SE	*p*
Intercept	415.72	0.79	<.001	415.71	1.04	<.001	415.60	1.42	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.02	.02	.136
Region: West	-.77	.84	.361	-1.42	1.03	.171	-2.79	1.29	.032
Region: Central East	.23	.88	.798	.19	1.08	.858	-.87	1.38	.528
Region: East	.64	.90	.480	1.72	1.09	.115	1.89	1.38	.172
Male	-..92	.40	.021	-.99	.58	.088	-.93	.84	.264
Black	-3.13	.72	<.001	-2.49	1.01	.014	-2.11	1.30	.264
Hispanic	-6.85	1.00	<.001	-8.72	1.56	<.001	-8.05	2.34	<.001
Native language other than English	-1.83	1.11	.099	-1.43	1.65	.385	-4.17	2.45	.090
Time between start of school and testing	.10	.02	<.001	.14	.03	<.001	.16	.04	<.001
Has an IEP	-7.95	.69	<.001	-8.55	1.00	<.001	-9.18	1.44	<.001
TN-VPK participation	7.60	.84	<.001	6.46	1.22	<.001	6.78	1.76	<.001
TN-VPK participation x Black	.96	.86	.261	2.08	1.27	.100	2.34	1.83	.201

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table H-1
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Hispanic and Non-Hispanic Children

Outcome and bandwidth	TN-VPK		No Pre-K		Pooled SD	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Composite								
12-month bandwidth					12.92			<.001
Hispanic	406.49	168	386.77	186		19.72	1.526	
Non-Hispanic	412.43	1910	402.28	1880		10.15	.786	
6-month bandwidth					12.66			<.001
Hispanic	405.13	78	384.82	87		20.31	1.604	
Non-Hispanic	412.63	931	402.66	971		9.97	.788	
3 -month bandwidth					11.94			.011
Hispanic	404.09	35	385.01	41		19.08	1.598	
Non-Hispanic	412.50	484	402.24	490		10.26	.859	
Letter-Word								
12-month bandwidth					21.15			
Hispanic	339.68	168	312.18	186		27.50	1.300	.038
Non-Hispanic	347.42	1910	325.89	1880		21.53	1.018	
6-month bandwidth					19.71			.028
Hispanic	339.93	78	309.28	87		30.65	1.555	
Non-Hispanic	347.48	931	326.36	971		21.12	1.072	
3 -month bandwidth					19.17			.151
Hispanic	341.01	35	312.90	41		28.11	1.466	
Non-Hispanic	345.87	484	326.62	490		19.25	1.004	

Appendix Table H-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Hispanic and Non-Hispanic Children

Outcome and bandwidth	TN-VPK		No Pre-K		Pooled SD	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Spelling								
12-month bandwidth					20.82			.192
Hispanic	376.90	168	354.58	186		23.32	1.072	
Non-Hispanic	379.77	1910	359.59	1880		20.18	.969	
6-month bandwidth					20.62			.725
Hispanic	376.06	78	354.83	87		21.23	1.030	
Non-Hispanic	380.03	931	360.91	971		19.12	.927	
3 -month bandwidth					20.43			.333
Hispanic	370.84	35	355.03	41		15.81	.774	
Non-Hispanic	379.87	484	359.84	490		20.03	.980	
Oral Comprehension								
12-month bandwidth					13.71			.068
Hispanic	443.15	168	436.93	186		6.22	.454	
Non-Hispanic	452.33	1910	449.35	1880		2.98	.217	
6-month bandwidth					13.08			.015
Hispanic	441.53	78	433.41	87		8.12	.621	
Non-Hispanic	452.15	931	449.66	971		2.49	.190	
3 -month bandwidth					12.87			.303
Hispanic	440.20	35	432.67	41		7.53	.585	
Non-Hispanic	453.11	484	449.33	490		3.78	.294	

Appendix Table H-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Hispanic and Non-Hispanic Children

Outcome and bandwidth	TN-VPK		No Pre-K		Pooled SD	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Picture Vocabulary								
12-month bandwidth					10.72			<.001
Hispanic	459.56	168	436.00	186		23.56	2.198	
Non-Hispanic	465.34	1910	463.41	1880		1.93	.180	
6-month bandwidth					10.89			<.001
Hispanic	457.79	78	433.09	87		24.70	2.268	
Non-Hispanic	465.61	931	463.64	971		1.97	.181	
3 -month bandwidth					10.41			<.001
Hispanic	460.06	35	435.02	41		25.04	2.405	
Non-Hispanic	465.10	484	464.03	490		1.07	.103	
Applied Problems								
12-month bandwidth					17.74			<.001
Hispanic	402.93	168	375.90	186		27.03	1.524	
Non-Hispanic	408.60	1910	401.88	1880		6.72	.379	
6-month bandwidth					18.12			<001
Hispanic	401.47	78	374.99	87		26.48	1.461	
Non-Hispanic	410.16	931	401.69	971		8.47	.467	
3 -month bandwidth					16.08			<.001
Hispanic	398.78	35	370.07	41		28.71	1.785	
Non-Hispanic	410.84	484	400.73	490		10.11	.629	

Appendix Table H-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Hispanic and Non-Hispanic Children

Outcome and bandwidth	TN-VPK		No Pre-K		Pooled SD	TN-VPK vs. No Pre-K difference	Effect size[d]	*p*[d]
	Mean[a]	*N*	Mean[a]	*N*				
Quantitative Concepts								
12-month bandwidth					13.99			.009
Hispanic	416.64	168	405.05	186		11.59	.828	
Non-Hispanic	421.12	1910	413.56	1880		7.56	.540	
6-month bandwidth					13.43			.043
Hispanic	414.02	78	403.07	87		10.95	.815	
Non-Hispanic	420.39	931	413.62	589		6.77	.504	
3 -month bandwidth					12.49			.371
Hispanic	413.51	35	404.07	41		9.44	.756	
Non-Hispanic	420.21	484	412.91	490		7.30	.584	

Note. Only children eligible for free or reduced price lunch and were either Hispanic or White were included. Previous analyses showed that the appropriate functional form was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Appendix Tables I-1-to I-7 report the full results for each model.

[b] Reported are the marginal means from the multilevel model.

[c] To facilitate comparison of these effect sizes with those reported for the overall pre-k effect for the free or reduced price lunch sample, effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the standard deviation of the treatment (TN-VPK) group overall.

[d] This is the *p*-value for the interaction term for condition (TN-VPK vs. No Pre-K) and whether the child was Hispanic or white. Estimates were based on a multilevel model with children nested within pre-k classroom. In addition to condition, Hispanic, and days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) whether the child was male; (3) whether the child was Black; (4) whether the child's native language was not English; (5) whether the child had an IEP placement; (6) the number of days elapsed between the Woodcock-Johnson testing date and the start of school (centered at the grand mean); and (7) the interaction term.

Appendix Table I-1
Statistical Models for the Effects of TN-VPK on the WJ Composite:
Comparison of Hispanic and Non-Hispanic Children

Effect	WJ Composite 12 months			WJ Composite 6 months			WJ Composite 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	406.13	.80	<.001	406.06	1.03	<.001	406.68	1.43	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.03	.02	.077
Region: West	-1.19	.88	.177	-1.55	1.05	.141	-2.54	1.36	.064
Region: Central East	.36	.91	.698	.55	1.10	.618	-.91	1.45	.530
Region: East	.03	.93	.975	.66	1.11	.552	.15	1.45	.919
Male	-2.31	.39	<.001	-1.70	.57	.003	-2.24	.82	.007
Black	-3.05	.58	<.001	-2.29	.80	.004	-2.07	1.07	.054
Hispanic	-15.51	1.14	<.001	-17.84	1.78	<.001	-17.23	2.57	<.001
Native language other than English	-7.44	1.11	<.001	-6.75	1.62	<.001	-9.22	2.43	<.001
Time between start of school and testing	.13	.02	<.001	.18	.03	<.001	.18	.04	<.001
Has an IEP	-10.52	.68	<.001	-11.00	.98	<.001	-11.68	1.42	<.001
TN-VPK participation	10.15	.80	<.001	9.97	1.15	<.001	10.25	1.67	<.001
TN-VPK participation x Hispanic	**9.57**	1.40	**<.001**	**10.34**	2.12	**<.001**	**8.83**	3.27	**.007**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. The composite is an average of the W-scores for the six subtests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table I-2
Statistical Models for the Effects of TN-VPK on the WJ Letter-Word Subtest: Comparison of Hispanic and Non-Hispanic Children

Effect	WJ Letter-Word 12 months			WJ Letter-Word 6 months			WJ Letter-Word 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	328.96	1.41	<.001	329.36	1.78	<.001	331.59	2.48	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.05	.03	.051
Region: West	-.38	1.57	.810	-1.49	1.85	.421	-2.69	2.43	.270
Region: Central East	-.56	1.64	.731	-1.08	1.93	.576	-2.67	2.58	.302
Region: East	-1.67	1.67	.320	-.45	1.94	.817	-1.41	2.59	.587
Male	-4.01	.68	<.001	-4.34	.96	<.001	-6.14	1.38	<.001
Black	1.46	1.02	.152	3.82	1.36	.005	3.41	1.84	.064
Hispanic	-13.71	1.98	<.001	-17.09	2.99	<.001	-13.72	4.33	.002
Native language other than English	-3.18	1.92	.097	-4.69	2.73	.086	-6.58	4.10	.109
Time between start of school and testing	.23	.04	<.001	.32	.05	<.001	.38	.07	<.001
Has an IEP	-10.20	1.18	<.001	-11.42	1.65	<.001	-11.71	2.38	<.001
TN-VPK participation	21.53	1.38	<.001	21.12	1.94	<.001	19.85	2.79	<.001
TN-VPK participation x Hispanic	**5.97**	2.41	**.013**	**9.54**	3.55	**.007**	8.86	5.48	.106

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table I-3
Statistical Models for the Effects of TN-VPK on the WJ Spelling Subtest: Comparison of Hispanic and Non-Hispanic Children

Effect	WJ Spelling 12 months			WJ Spelling 6 months			WJ Spelling 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	362.76	1.20	<.001	362.97	1.67	<.001	362.73	2.35	<.001
Days from age eligibility cutoff	.06	<.01	<.001	.06	.01	<.001	.04	.03	.134
Region: West	1.60	1.28	.211	.64	1.69	.704	-1.42	2.23	.527
Region: Central East	.66	1.33	.621	1.71	1.76	.333	-.58	2.38	.809
Region: East	2.35	1.36	.086	4.16	1.78	.020	3.27	2.38	.172
Male	-5.97	.64	<.001	-4.93	.93	<.001	-5.35	1.36	<.001
Black	-1.57	.92	.086	-.01	1.29	.991	1.86	1.77	.293
Hispanic	-5.01	1.85	.007	-6.07	2.89	.035	-4.81	4.25	.258
Native language other than English	.26	1.78	.882	.94	2.63	.722	-.49	4.02	.903
Time between start of school and testing	.22	.04	<.001	.27	.05	<.001	.31	.07	<.001
Has an IEP	-11.32	1.10	<.001	-12.36	1.60	<.001	-12.80	2.35	<.001
TN-VPK participation	20.18	1.29	<.001	19.12	1.88	<.001	20.03	2.77	<.001
TN-VPK participation x Hispanic	2.14	2.26	.345	2.11	3.44	.540	-4.12	5.41	.446

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table I-4
Statistical Models for the Effects of TN-VPK on the WJ Oral Comprehension Subtest: Comparison of Hispanic and Non-Hispanic Children

Effect	WJ Oral Comprehension 12 months			WJ Oral Comprehension 6 months			WJ Oral Comprehension 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	454.64	0.83	<.001	454.25	1.00	<.001	454.39	1.40	<.001
Days from age eligibility cutoff	.03	<.01	<.001	.03	.01	<.001	.01	.01	.413
Region: West	-3.34	.93	<.001	-3.01	1.16	.010	-3.09	1.36	.025
Region: Central East	1.38	.97	.157	1.86	1.21	.126	.62	1.44	.670
Region: East	-.84	.99	.396	-.84	1.22	.492	-.70	1.45	.628
Male	-1.52	.39	<.001	-.58	.55	.297	-.44	.79	.574
Black	-5.45	.60	<.001	-5.23	.81	<.001	-5.70	1.04	<.001
Hispanic	-12.42	1.15	<.001	-16.25	1.74	<.001	-16.67	2.47	<.001
Native language other than English	-12.60	1.11	<.001	-11.38	1.59	<.001	-12.42	2.34	<.001
Time between start of school and testing	.05	.02	.028	.09	.03	.005	.07	.04	.109
Has an IEP	-10.01	.68	<.001	-9.55	.94	<.001	-10.65	1.36	<.001
TN-VPK participation	2.99	.79	<.001	2.29	1.12	.027	3.78	1.61	.019
TN-VPK participation x Hispanic	**3.24**	1.40	**.021**	**5.62**	2.06	**.006**	3.75	3.13	.231

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table I-5
Statistical Models for the Effects of TN-VPK on the WJ Picture Vocabulary Subtest: Comparison of Hispanic and Non-Hispanic Children

Effect	WJ Picture Vocabulary 12 months			WJ Picture Vocabulary 6 months			WJ Picture Vocabulary 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	466.09	.79	<.001	466.11	1.02	<.001	467.71	1.40	<.001
Days from age eligibility cutoff	.02	<.01	<.001	.02	.01	<.001	.02	.02	.145
Region: West	-1.53	.87	.082	-1.26	1.05	.233	-1.93	1.34	.152
Region: Central East	.08	.91	.931	-.06	1.10	.959	-1.76	1.43	.220
Region: East	-.33	.93	.725	-.45	1.10	.685	-1.26	1.43	.381
Male	.48	.39	.211	1.30	.56	.020	1.34	.80	.096
Black	-1.77	.58	.002	-2.41	.78	.002	-3.03	1.05	.004
Hispanic	-27.41	1.12	<.001	-30.55	1.73	<.001	-29.01	2.52	<.001
Native language other than English	-15.52	1.09	<.001	-13.79	1.58	<.001	-18.78	2.38	<.001
Time between start of school and testing	.03	.02	.255	.04	.03	.164	.02	.04	.664
Has an IEP	-7.83	.67	<.001	-7.44	.96	<.001	-7.46	1.39	<.001
TN-VPK participation	1.93	.78	.014	1.98	1.12	.079	1.07	1.64	.514
TN-VPK participation x Hispanic	**21.63**	1.37	**<.001**	**22.72**	2.06	**<.001**	**23.97**	3.20	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table I-6
Statistical Models for the Effects of TN-VPK on the WJ Applied Problems Subtest: Comparison of Hispanic and Non-Hispanic Children

Effect	WJ Applied Problems 12 months			WJ Applied Problems 6 months			WJ Applied Problems 3 months		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	408.50	1.25	<.001	408.05	1.55	<.001	408.46	2.15	<.001
Days from age eligibility cutoff	.05	<.01	<.001	.04	.01	<.001	.02	.02	.446
Region: West	-2.72	1.36	.047	-2.94	1.51	.054	-3.44	2.05	.096
Region: Central East	.46	1.42	.747	.75	1.58	.635	-.20	2.19	.927
Region: East	.20	1.44	.890	-.16	1.59	.920	-1.02	2.19	.643
Male	-1.98	.63	.002	-.59	.89	.509	-1.96	1.24	.113
Black	-8.08	.93	<.001	-8.17	1.20	<.001	-7.95	1.61	<.001
Hispanic	-25.98	1.84	<.001	-26.70	2.76	<.001	-30.65	3.87	<.001
Native language other than English	-11.48	1.78	<.001	-10.12	2.51	<.001	-13.00	3.66	<.001
Time between start of school and testing	.13	.04	<.001	.20	.05	<.001	.14	.06	.029
Has an IEP	-15.85	1.10	<.001	-16.77	1.53	<.001	-18.52	2.14	<.001
TN-VPK participation	6.72	1.28	<.001	8.47	1.80	<.001	10.11	2.52	<.001
TN-VPK participation x Hispanic	**20.31**	2.25	**<.001**	**18.01**	3.30	**<.001**	**18.59**	4.92	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table I-7
Statistical Models for the Effects of TN-VPK on the WJ Quantitative Concepts Subtest: Comparison of Hispanic and Non-Hispanic Children

Effect	WJ Quantitative Concepts 12 months			WJ Quantitative Concepts 6 months			WJ Quantitative Concepts 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	415.70	0.78	<.001	415.56	1.03	<.001	415.29	1.39	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.02	.02	.132
Region: West	-.71	.85	.402	-1.37	1.03	.185	-2.74	1.28	.035
Region: Central East	.29	.88	.741	.25	1.07	.816	-.82	1.37	.551
Region: East	.67	.90	.457	1.75	1.08	.108	1.95	1.37	.158
Male	-.90	.40	.023	-1.00	.58	.085	-.85	.83	.310
Black	-2.66	.58	<.001	-1.49	.80	.063	-.92	1.05	.382
Hispanic	-8.51	1.16	<.001	-10.55	1.80	<.001	-8.83	2.59	<.001
Native language other than English	-2.12	1.11	.057	-1.33	1.64	.421	-4.01	2.45	.102
Time between start of school and testing	.10	.02	<.001	.14	.03	<.001	.16	.04	<.001
Has an IEP	-7.93	.69	<.001	-8.53	1.00	<.001	-9.15	1.44	<.001
TN-VPK participation	7.55	.81	<.001	6.77	1.17	<.001	7.31	1.69	<.001
TN-VPK participation x Hispanic	**4.03**	1.41	**.004**	4.18	2.15	.052	2.13	3.32	.522

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table J-1
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Children Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Composite								
12-month bandwidth					12.92			<.001
English	411.96	1914	401.89	1926		10.07	.779	
Another language	410.12	164	389.00	140		21.12	1.635	
6-month bandwidth					12.66			<.001
English	412.04	936	402.14	980		9.90	.782	
Another language	411.30	73	390.28	78		21.02	1.660	
3 -month bandwidth					11.94			.003
English	411.98	487	401.97	492		10.01	.838	
Another language	409.09	32	388.96	39		20.13	1.686	
Letter-Word								
12-month bandwidth					21.15			<.001
English	346.55	1914	325.33	1926		21.22	1.003	
Another language	348.35	164	316.85	140		31.50	1.489	
6-month bandwidth					19.71			.028
English	346.67	936	325.82	980		20.85	1.058	
Another language	348.92	73	315.18	78		33.74	1.712	
3 -month bandwidth					19.17			.016
English	345.27	487	326.58	492		18.69	.975	
Another language	347.17	32	314.80	39		32.37	1.689	

Appendix Table J-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Children Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Spelling								
12-month bandwidth					20.82			.006
English	379.24	1914	359.40	1926		19.84	.953	
Another language	382.60	164	356.12	140		26.48	1.272	
6-month bandwidth					20.62			.044
English	379.36	936	360.63	980		18.73	.908	
Another language	384.19	73	358.21	78		25.98	1.260	
3 -month bandwidth					20.43			.944
English	379.29	487	359.57	492		19.72	.965	
Another language	378.95	32	358.83	39		20.12	.985	
Oral Comprehension								
12-month bandwidth					13.71			.213
English	452.37	1914	449.27	1926		3.10	.226	
Another language	440.85	164	435.86	140		4.99	.264	
6-month bandwidth					13.08			.132
English	451.98	936	449.31	980		2.67	.204	
Another language	442.37	73	436.46	78		5.91	.452	
3 -month bandwidth					12.87			.807
English	452.94	487	448.98	492		3.96	.308	
Another language	441.08	32	436.32	39		4.76	.370	

Appendix Table J-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Children Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Picture Vocabulary								
12-month bandwidth					10.72			<.001
English	465.04	1914	462.85	1926		2.20	.205	
Another language	459.72	164	438.14	140		21.58	2.013	
6-month bandwidth					10.89			<.001
English	465.22	936	462.82	980		2.40	.220	
Another language	460.68	73	441.32	78		19.36	1.778	
3 -month bandwidth					10.41			<.001
English	464.84	487	463.89	492		.95	.091	
Another language	459.26	32	437.42	39		21.84	2.098	
Applied Problems								
12-month bandwidth					17.74			<.001
English	407.95	1914	401.26	1926		6.69	.377	
Another language	407.54	164	379.24	140		28.30	1.595	
6-month bandwidth					18.12			<001
English	409.34	936	401.02	980		8.32	.459	
Another language	409.88	73	381.93	78		27.95	1.542	
3 -month bandwidth					16.08			<.001
English	409.79	487	400.08	492		9.71	.604	
Another language	409.20	32	379.66	39		29.54	1.837	

Appendix Table J-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Composite and Subtests: Comparison of Children Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*P*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Quantitative Concepts								
12-month bandwidth					13.99			.001
English	420.61	1914	413.24	1926		7.37	..527	
Another language	421.68	164	407.74	140		13.94	.996	
6-month bandwidth					13.43			.003
English	419.69	936	413.13	980		6.56	.488	
Another language	422.02	73	408.68	78		13.34	.993	
3 -month bandwidth					12.49			.123
English	419.76	487	412.74	492		7.02	.562	
Another language	419.01	32	406.69	39		12.32	.986	

Note. Only children eligible for free or reduced price lunch were included. Previous analyses showed that the appropriate functional form was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Appendix Tables K-1 to K-7 report the full results for each model.

[a] Reported are the marginal means from the multilevel model.

[b] To facilitate comparison of these effect sizes with those reported for the overall pre-k effect, effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the standard deviation of the treatment (TN-VPK) group for the sample of children eligible for free or reduced price lunch.

[c] This is the *p*-value for the interaction term for condition (TN-VPK vs. No Pre-K) and whether the child's native language was English or another language. Estimates were based on a multilevel model with children nested within pre-k classroom. In addition to condition, native language, and days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) whether the child was male; (3) whether the child was Black; (4) whether the child was Hispanic; (5) whether the child had an IEP placement; (6) the number of days elapsed between the Woodcock-Johnson testing date and the start of school (centered at the grand mean); and (7) the interaction term

Appendix Table K-1
Statistical Models for the Effects of TN-VPK on the WJ Composite:
Comparison of Children Whose Primary Language was English or Another Language

Effect	WJ Composite 12 months			WJ Composite 6 months			WJ Composite 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	406.34	.80	<.001	406.36	1.04	<.001	406.93	1.43	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.03	.02	.057
Region: West	-1.56	.88	.077	-2.04	1.06	.057	-2.83	1.36	.039
Region: Central East	.22	.92	.814	.40	1.11	.722	-.89	1.45	.537
Region: East	-.13	.93	.891	.45	1.12	.686	.08	1.45	.957
Male	-2.32	.39	<.001	-1.71	.57	.003	-2.22	.82	.007
Black	-3.02	.59	<.001	-2.26	.80	.005	-2.06	1.07	.054
Hispanic	-11.33	.99	<.001	-13.36	1.53	<.001	-11.10	2.32	<.001
Native language other than English	-12.90	1.39	<.001	-11.86	1.92	<.001	-13.01	2.77	<.001
Time between start of school and testing	.13	.02	<.001	.18	.03	<.001	.18	.04	<.001
Has an IEP	-10.54	.68	<.001	-11.04	.98	<.001	-11.80	1.42	<.001
TN-VPK participation	10.15	.80	<.001	9.90	1.16	<.001	10.00	1.68	<.001
TN-VPK participation x Non-English	**11.06**	1.50	**<.001**	**11.12**	2.22	**<.001**	**10.13**	3.38	**.003**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. The composite is an average of the W-scores for the six subtests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table K-2
Statistical Models for the Effects of TN-VPK on the WJ Letter-Word Subtest: Comparison of Children Whose Primary Language was English or Another Language

Effect	WJ Letter-Word 12 months			WJ Letter-Word 6 months			WJ Letter-Word 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	329.23	1.42	<.001	329.75	1.79	<.001	332.01	2.49	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.05	.03	.039
Region: West	-.68	1.57	.666	-2.01	1.85	.279	-3.09	2.43	.207
Region: Central East	-.64	1.64	.697	-1.23	1.93	.525	-2.65	2.58	.306
Region: East	-1.78	1.66	.286	-.68	1.95	.726	-1.54	2.59	.554
Male	-4.01	.68	<.001	-4.35	.96	<.001	-6.11	1.37	<.001
Black	1.49	1.02	.146	3.84	1.36	.005	3.44	1.84	.061
Hispanic	-11.03	1.71	<.001	-12.93	2.57	<.001	-10.52	3.90	.007
Native language other than English	-8.48	2.40	<.001	-10.64	3.23	.001	-11.78	4.66	.012
Time between start of school and testing	.24	.04	<.001	.32	.05	<.001	.37	.07	<.001
Has an IEP	-10.21	1.18	<.001	-11.44	1.65	<.001	-11.87	2.38	<.001
TN-VPK participation	21.22	1.38	<.001	20.85	1.94	<.001	18.70	2.83	<.001
TN-VPK participation x Non-English	**10.28**	2.60	**<.001**	**12.89**	3.71	**<.001**	**13.67**	5.65	**.016**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table K-3
Statistical Models for the Effects of TN-VPK on the WJ Spelling Subtest: Comparison of Children Whose Primary Language was English or Another Language

Effect	WJ Spelling 12 months			WJ Spelling 6 months			WJ Spelling 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	362.99	1.21	<.001	363.28	1.67	<.001	362.86	2.37	<.001
Days from age eligibility cutoff	.06	<.01	<.001	.06	.01	<.001	.04	.03	.133
Region: West	1.44	1.27	.261	.39	1.69	.815	-1.42	2.24	.528
Region: Central East	.64	1.33	.630	1.67	1.76	.344	-.57	2.38	.811
Region: East	2.30	1.36	.093	4.06	1.77	.024	3.22	2.39	.180
Male	-5.97	.63	<.001	-4.95	.93	<.001	-5.34	1.36	<.001
Black	-1.56	.91	.088	-.01	1.29	.996	1.87	1.77	.291
Hispanic	-3.99	1.60	.013	-5.12	2.49	.040	-6.20	3.84	.106
Native language other than English	-3.28	2.23	.142	-2.42	3.12	.438	-.74	4.58	.872
Time between start of school and testing	.22	.04	<.001	.27	.05	<.001	.31	.07	<.001
Has an IEP	-11.31	1.10	<.001	-12.35	1.60	<.001	-12.79	2.35	<.001
TN-VPK participation	19.83	1.29	<.001	18.73	1.88	<.001	19.72	2.79	<.001
TN-VPK participation x Non-English	**6.64**	2.44	**.006**	**7.26**	3.60	**.044**	.40	5.60	.944

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table K-4
Statistical Models for the Effects of TN-VPK on the WJ Oral Comprehension Subtest: Comparison of Children Whose Primary Language was English or Another Language

Effect	WJ Oral Comprehension 12 months			WJ Oral Comprehension 6 months			WJ Oral Comprehension 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	454.62	0.84	<.001	454.25	1.09	<.001	454.32	1.41	<.001
Days from age eligibility cutoff	.03	<.01	<.001	.03	.01	<.001	.01	.02	.408
Region: West	-3.43	.93	<.001	-3.11	1.16	.008	-3.13	1.36	.023
Region: Central East	1.33	.97	.173	1.79	1.21	.141	.61	1.44	.672
Region: East	-.88	.99	.373	-.91	1.22	.455	-.67	1.45	.642
Male	-1.53	.39	<.001	-.58	.56	.298	-.44	.79	.574
Black	-5.45	.60	<.001	-5.21	.81	<.001	-5.71	1.04	<.001
Hispanic	-11.05	1.00	<.001	-13.83	1.50	<.001	-15.38	2.23	<.001
Native language other than English	-13.40	1.39	<.001	-12.85	1.89	<.001	-12.65	2.66	<.001
Time between start of school and testing	.05	.02	.028	.09	.03	.005	.07	.04	.103
Has an IEP	-10.02	.68	<.001	-9.57	.96	<.001	-10.67	1.36	<.001
TN-VPK participation	3.10	.80	<.001	2.66	1.13	.018	3.96	1.62	.014
TN-VPK participation x Non-English	1.88	1.51	.213	3.25	2.16	.132	.79	3.24	.807

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table K-5
Statistical Models for the Effects of TN-VPK on the WJ Picture Vocabulary Subtest: Comparison of Children Whose Primary Language was English or Another Language

Effect	WJ Picture Vocabulary 12 months			WJ Picture Vocabulary 6 months			WJ Picture Vocabulary 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	466.25	.82	<.001	466.27	1.06	<.001	467.99	1.43	<.001
Days from age eligibility cutoff	.02	<.01	<.001	.02	.01	<.001	.03	.02	.084
Region: West	-2.29	.91	.014	-2.11	1.10	.056	-2.57	1.36	.062
Region: Central East	-.23	.95	.809	-.35	1.15	.761	-1.73	1.45	.236
Region: East	-.62	.97	.525	-.75	1.15	.514	-1.32	1.45	.365
Male	.47	.39	.228	1.27	.56	.023	1.37	.81	.092
Black	-1.62	.59	.006	-2.23	.80	.006	-2.98	1.06	.005
Hispanic	-18.04	.98	<.001	-20.71	1.52	<.001	-20.59	2.29	<.001
Native language other than English	-24.71	1.38	<.001	-21.50	1.90	<.001	-26.47	2.73	<.001
Time between start of school and testing	.03	.02	.249	.04	.03	.172	.02	.04	.610
Has an IEP	-7.89	.68	<.001	-7.56	.97	<.001	-7.75	1.40	<.001
TN-VPK participation	2.19	.79	.006	2.40	1.14	.036	.95	1.66	.568
TN-VPK participation x Non-English	**19.38**	1.49	**<.001**	**16.96**	2.19	**<.001**	**20.89**	3.33	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table K-6
Statistical Models for the Effects of TN-VPK on the WJ Applied Problems Subtest: Comparison of Children Whose Primary Language was English or Another Language

Effect	WJ Applied Problems 12 months			WJ Applied Problems 6 months			WJ Applied Problems 3 months		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	408.85	1.27	<.001	408.54	1.56	<.001	408.88	2.17	<.001
Days from age eligibility cutoff	.05	<.01	<.001	.04	.01	<.001	.02	.02	.347
Region: West	-3.48	1.38	.013	-3.79	1.54	.015	-4.04	2.06	.052
Region: Central East	.14	1.44	.920	.48	1.60	.764	-.17	2.19	.938
Region: East	-.12	1.47	.937	-.51	1.61	.754	-1.14	2.20	.606
Male	-1.99	.63	.002	-.62	.89	.486	-1.93	1.23	.119
Black	-8.01	.94	<.001	-8.11	1.21	<.001	-7.92	1.61	<.001
Hispanic	-17.14	1.60	<.001	-18.89	2.38	<.001	-24.06	3.49	<.001
Native language other than English	-22.02	2.23	<.001	-19.09	2.98	<.001	-20.41	4.17	<.001
Time between start of school and testing	.13	.04	<.001	.20	.05	<.001	.14	.06	.027
Has an IEP	-15.90	1.10	<.001	-16.85	1.53	<.001	-18.77	2.14	<.001
TN-VPK participation	6.69	1.28	<.001	8.31	1.80	<.001	9.71	2.53	<.001
TN-VPK participation x Non-English	**22.61**	2.42	**<.001**	**19.64**	3.46	**<.001**	**19.83**	5.09	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table K-7
Statistical Models for the Effects of TN-VPK on the WJ Quantitative Concepts Subtest: Comparison of Children Whose Primary Language was English or Another Language

Effect	WJ Quantitative Concepts 12 months			WJ Quantitative Concepts 6 months			WJ Quantitative Concepts 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	415.87	0.78	<.001	415.79	1.04	<.001	415.49	1.40	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.02	.02	.116
Region: West	-.91	.85	.286	-1.63	1.03	.116	-2.89	1.29	.027
Region: Central East	.24	.88	.789	.18	1.08	.868	-.81	1.38	.555
Region: East	.59	.90	.512	1.63	1.09	.135	1.88	1.37	.173
Male	-.90	.40	.022	-1.01	.58	.081	-.83	.83	.316
Black	-2.65	.58	<.001	-1.48	.80	.064	-.92	1.05	.382
Hispanic	-6.71	1.00	<.001	-8.73	1.56	<.001	-8.04	2.33	<.001
Native language other than English	-5.50	1.40	.057	-4.45	1.95	.022	-6.05	2.79	.031
Time between start of school and testing	.10	.02	<.001	.15	.03	<.001	.16	.04	<.001
Has an IEP	-7.93	.69	<.001	-8.54	1.00	<.001	-9.21	1.44	<.001
TN-VPK participation	7.37	.80	<.001	6.56	1.17	<.001	7.02	1.70	<.001
TN-VPK participation x Non-English	**6.57**	1.52	**<.001**	**6.78**	2.25	**.003**	5.30	3.43	.123

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table L-1
Estimated Impact of the TN-VPK Program on the WJ Subtests: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Letter-Word								
12-month bandwidth					21.15			<.001
Hispanic and English	342.63	49	313.85	72		28.78	1.361	
Hispanic and another language	335.37	119	308.04	114		27.33	1.292	
Not Hispanic and English	347.37	1865	326.29	1854		21.08	.997	
Not Hispanic and another language	357.78	45	310.90	26		46.88	2.217	
6-month bandwidth					19.71			.043
Hispanic and English	340.08	25	310.35	27		29.73	1.508	
Hispanic and another language	335.19	53	304.33	60		30.86	1.565	
Not Hispanic and English	347.59	911	326.94	953		20.65	1.048	
Not Hispanic and another language	355.37	20	311.89	18		43.48	2.206	
Spelling								
12-month bandwidth					20.82			.202
Hispanic and English	376.25	49	357.41	72		18.84	.905	
Hispanic and another language	377.10	119	352.84	114		24.26	1.165	
Not Hispanic and English	379.56	1865	359.68	1854		19.88	.955	
Not Hispanic and another language	388.44	45	354.17	26		34.27	1.646	
6-month bandwidth					20.62			<.001
Hispanic and English	374.88	25	355.99	27		18.89	.916	
Hispanic and another language	376.99	53	354.97	60		22.02	1.068	
Not Hispanic and English	379.74	911	361.06	953		18.68	.906	
Not Hispanic and another language	392.03	20	352.44	18		39.59	1.920	

Appendix Table L-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Subtests: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Oral Comprehension								
12-month bandwidth					13.71			<.001
Hispanic and English	447.48	49	435.13	72		12.35	.901	
Hispanic and another language	430.20	119	426.93	114		3.27	.239	
Not Hispanic and English	453.21	1865	450.33	1854		2.88	.210	
Not Hispanic and another language	443.26	45	431.25	26		12.01	.876	
6-month bandwidth					13.08			<.001
Hispanic and English	449.21	25	427.74	27		21.47	1.641	
Hispanic and another language	428.03	53	425.46	60		2.57	.196	
Not Hispanic and English	452.87	911	450.62	953		2.25	.172	
Not Hispanic and another language	448.04	20	431.09	18		16.95	1.296	
Picture Vocabulary								
12-month bandwidth					10.72			<.001
Hispanic and English	461.58	49	434.91	72		26.67	2.488	
Hispanic and another language	444.62	119	422.94	114		21.68	2.022	
Not Hispanic and English	466.35	1865	464.75	1854		1.60	.149	
Not Hispanic and another language	456.84	45	434.40	26		22.44	2.093	
6-month bandwidth					10.89			<.001
Hispanic and English	463.27	25	424.01	27		39.26	3.605	
Hispanic and another language	443.04	53	424.64	60		18.40	1.690	
Not Hispanic and English	466.52	911	464.93	953		1.59	.146	
Not Hispanic and another language	457.55	20	435.88	18		21.67	1.990	

Appendix Table L-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Subtests: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Applied Problems								
12-month bandwidth					17.74			<.001
Hispanic and English	404.99	49	376.91	72		28.08	1.583	
Hispanic and another language	391.46	119	364.93	114		26.53	1.495	
Not Hispanic and English	409.17	1865	402.97	1854		6.20	.349	
Not Hispanic and another language	410.17	45	374.48	26		35.69	2.012	
6-month bandwidth					18.12			<.001
Hispanic and English	403.86	25	371.66	27		32.20	1.777	
Hispanic and another language	391.00	53	367.20	60		23.70	1.308	
Not Hispanic and English	410.60	911	402.82	953		7.78	.429	
Not Hispanic and another language	415.96	20	374.14	18		41.82	2.308	
Quantitative Concepts								
12-month bandwidth					13.99			<.001
Hispanic and English	418.04	49	405.75	72		12.29	.878	
Hispanic and another language	414.00	119	402.59	114		11.41	.928	
Not Hispanic and English	421.11	1865	413.84	1854		7.27	.520	
Not Hispanic and another language	426.82	45	403.27	26		23.55	1.916	

Appendix Table L-1 (*continued*)
Estimated Impact of the TN-VPK Program on the WJ Subtests: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Outcome and bandwidth	TN-VPK		No Pre-K		TN-VPK Standard Deviation	TN-VPK vs. No Pre-K difference	Effect size[b]	*p*[c]
	Mean[a]	*N*	Mean[a]	*N*				
Quantitative Concepts								
6-month bandwidth					13.43			<.001
Hispanic and English	415.64	25	401.53	27		14.11	1.075	
Hispanic and another language	411.83	53	402.36	60		9.47	.705	
Not Hispanic and English	420.30	911	413.93	953		6.37	.474	
Not Hispanic and another language	429.11	20	402.58	18		26.53	1.975	

Note. Only children eligible for free or reduced price lunch were included. Previous analyses showed that the appropriate functional form was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Appendix Tables M-1-to M-7 report the full results for each model.

[a] Reported are the marginal means from the multilevel model.

[b] To facilitate comparison of these effect sizes with those reported for the overall pre-k effect, effect sizes were calculated by dividing the TN-VPK vs. No Pre-K difference by the standard deviation of the treatment (TN-VPK) group for the sample of children eligible for free or reduced price lunch.

[c] This is the *p*-value for the interaction term for condition (TN-VPK vs. No Pre-K), whether the child was Hispanic, and whether the child's native language was English or another language. Estimates were based on a multilevel model with children nested within pre-k classroom. In addition to condition, Hispanic, Language, and days from the eligibility cutoff date (centered at zero), covariates included: (1) Region; (2) whether the child was male; (3) whether the child was Black; (4) whether the child had an IEP placement; (5) the number of days elapsed between the Woodcock-Johnson testing date and the start of school (centered at the grand mean); and (6) the interaction term

Appendix Table M-1
Statistical Models for the Effects of TN-VPK on the WJ Composite: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Effect	WJ Composite 12 months			WJ Composite 6 months			WJ Composite 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	406.44	0.80	<.001	406.51	1.04	<.001	406.95	1.44	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.03	.02	.064
Region: West	-1.38	.88	.120	-1.67	1.06	.117	-2.62	1.37	.058
Region: Central East	.25	.92	.783	.52	1.10	.640	-.80	1.45	.582
Region: East	-.11	.93	.906	.44	1.11	.695	.08	1.45	.959
Male	-2.32	.39	<.001	-1.78	.57	.002	-2.25	.82	.006
Black	-3.16	.58	<.001	-2.49	.79	.002	-2.62	1.07	.044
Hispanic	-15.67	1.47	<.001	-21.49	2.47	<.001	-18.05	3.49	<.001
Native language other than English	-18.27	2.54	<.001	-18.79	3.10	<.001	-16.43	4.48	<.001
Time between start of school and testing	.13	.02	<.001	.18	.03	<.001	.18	.04	<.001
Has an IEP	-10.52	.68	<.001	-10.98	.98	<.001	-11.75	1.42	<.001
TN-VPK participation	9.81	.80	<.001	9.53	1.15	<.001	9.92	1.69	<.001
TN-VPK participation x Hispanic	11.37	2.43	<.001	16.35	3.81	<.001	6.07	6.14	.323
TN-VPK participation x Non-English	19.31	3.23	<.001	22.16	4.52	<.001	11.60	6.97	.096
Hispanic x Non-English	10.69	3.11	<.001	16.73	4.21	<.001	8.22	6.08	.177
TN-VPK participation x Hispanic x Non-English	**-21.40**	4.32	**<.001**	**-30.23**	6.34	**<.001**	-7.69	9.96	.440

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. The composite is an average of the W-scores for the six subtests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom, and Hispanic children whose primary language was not English and who were in the treatment group served as the referent group for the parameter estimate. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table M-2
Statistical Models for the Effects of TN-VPK on the WJ Letter-Word Subtest: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Effect	WJ Letter-Word 12 months			WJ Letter-Word 6 months		
	b	SE	*p*	*b*	SE	*p*
Intercept	329.37	1.42	<.001	329.85	1.80	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001
Region: West	-.69	1.58	.662	-1.88	1.87	.314
Region: Central East	-.73	1.64	.659	-1.21	1.94	.533
Region: East	-1.88	1.68	.263	-.76	1.95	.697
Male	-4.03	.68	<.001	-4.41	.96	<.001
Black	-1.31	1.02	.200	3.67	1.36	.007
Hispanic	-12.45	2.55	<.001	-16.59	4.18	<.001
Native language other than English	-15.39	4.41	<.001	-15.05	5.24	.004
Time between start of school and testing	.24	.04	<.001	.32	.05	<.001
Has an IEP	-10.19	1.18	<.001	-11.42	1.65	<.001
TN-VPK participation	21.07	1.38	<.001	20.65	1.94	<.001
TN-VPK participation x Hispanic	7.71	4.21	.067	9.08	6.44	.159
TN-VPK participation x Non-English	25.80	5.59	<.001	22.283	7.63	.003
Hispanic x Non-English	9.58	5.39	.076	9.02	7.11	.205
TN-VPK participation x Hispanic x Non-English	**-27.25**	7.49	**<.001**	**-21.70**	10.69	**.043**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom, and Hispanic children whose primary language was not English and who were in the treatment group served as the referent group for the parameter estimate. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table M-3
Statistical Models for the Effects of TN-VPK on the WJ Spelling Subtest: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Effect	WJ Spelling 12 months			WJ Spelling 6 months		
	b	SE	*p*	*b*	SE	*p*
Intercept	363.02	1.20	<.001	363.43	1.68	<.001
Days from age eligibility cutoff	.06	<.01	<.001	.06	.01	<.001
Region: West	1.30	1.27	.306	.24	1.70	.890
Region: Central East	.56	1.32	.673	1.60	1.76	.366
Region: East	2.24	1.35	.100	3.88	1.78	.030
Male	-5.99	.63	<.001	-5.00	.93	<.001
Black	-1.60	.91	.081	-.12	1.29	.926
Hispanic	-2.27	2.38	.340	-5.07	4.04	.209
Native language other than English	-5.51	4.13	.182	-8.62	5.06	.089
Time between start of school and testing	.22	.04	<.001	.27	.05	<.001
Has an IEP	-11.31	1.20	<.001	-12.36	1.60	<.001
TN-VPK participation	19.88	1.29	<.001	18.68	1.88	<.001
TN-VPK participation x Hispanic	-1.04	3.94	.793	.21	6.23	.973
TN-VPK participation x Non-English	14.39	5.25	.006	20.91	7.39	.005
Hispanic x Non-English	.94	5.06	.853	7.60	6.88	.269
TN-VPK participation x Hispanic x Non-English	-8.96	7.02	.202	-17.79	10.35	.086

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom, and Hispanic children whose primary language was not English and who were in the treatment group served as the referent group for the parameter estimate. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table M-4
Statistical Models for the Effects of TN-VPK on the WJ Oral Comprehension Subtest: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Effect	WJ Oral Comprehension 12 months			WJ Oral Comprehension 6 months		
	b	SE	*p*	*b*	SE	*p*
Intercept	454.74	.83	<.001	454.44	1.08	<.001
Days from age eligibility cutoff	.03	<.01	<.001	.03	.01	<.001
Region: West	-3.25	.93	<.001	-2.67	1.15	.022
Region: Central East	1.37	.97	.161	1.90	1.20	.115
Region: East	-.87	.99	.376	-.96	1.21	.428
Male	-1.53	.39	<.001	-.66	.55	.232
Black	-5.58	..60	<.001	-5.52	.81	<.001
Hispanic	-15.20	1.48	<.001	-22.88	2.42	<.001
Native language other than English	-19.08	2.56	<.001	-19.53	3.03	<.001
Time between start of school and testing	.06	.02	.024	.09	.03	.004
Has an IEP	-10.00	.68	<.001	-9.51	.95	<.001
TN-VPK participation	2.88	.80	<.001	2.24	1.12	.045
TN-VPK participation x Hispanic	9.47	2.44	<.001	19.23	3.72	<.001
TN-VPK participation x Non-English	9.12	3.25	.005	14.70	4.40	<.001
Hispanic x Non-English	10.88	3.13	<.001	17.25	4.11	<.001
TN-VPK participation x Hispanic x Non-English	**-18.20**	4.35	**<.001**	**-33.61**	6.17	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom, and Hispanic children whose primary language was not English and who were in the treatment group served as the referent group for the parameter estimate. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table M-5
Statistical Models for the Effects of TN-VPK on the WJ Picture Vocabulary Subtest: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Effect	WJ Picture Vocabulary 12 months			WJ Picture Vocabulary 6 months		
	b	SE	*p*	*b*	SE	*p*
Intercept	466.43	.79	<.001	466.43	1.02	<.001
Days from age eligibility cutoff	.02	<.01	<.001	.02	.01	<.001
Region: West	-1.62	.88	.066	-.94	1.05	.374
Region: Central East	<.01	.91	.998	.05	1.09	.961
Region: East	-.46	.93	.624	-.54	1.10	.622
Male	.48	.38	.213	1.21	.55	.028
Black	-1.92	.58	<.001	-2.68	.77	<.001
Hispanic	-29.84	1.45	<.001	-40.92	2.39	<.001
Native language other than English	-30.36	2.50	<.001	-29.06	3.00	<.001
Time between start of school and testing	.03	.02	.213	.05	.03	.120
Has an IEP	-7.82	.67	<.001	-7.38	.94	<.001
TN-VPK participation	1.60	.78	.041	1.58	1.11	.156
TN-VPK participation x Hispanic	25.07	2.39	<.001	37.67	3.69	<.001
TN-VPK participation x Non-English	20.85	3.17	<.001	20.09	4.37	<.001
Hispanic x Non-English	18.38	3.06	<.001	29.68	4.07	<.001
TN-VPK participation x Hispanic x Non-English	**-25.83**	4.24	**<.001**	**-40.94**	6.12	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom, and Hispanic children whose primary language was not English and who were in the treatment group served as the referent group for the parameter estimate. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table M-6
Statistical Models for the Effects of TN-VPK on the WJ Applied Problems Subtest: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Effect	WJ Applied Problems 12 months			WJ Applied Problems 6 months		
	b	SE	*p*	*b*	SE	*p*
Intercept	408.98	1.26	<.001	408.73	1.55	<.001
Days from age eligibility cutoff	.05	<.01	<.001	.04	.01	<.001
Region: West	-3.02	1.36	.028	-3.22	1.53	..037
Region: Central East	.30	1.42	.833	.70	1.59	.662
Region: East	-.01	1.45	.995	-.49	1.60	.761
Male	-2.00	.63	.002	-.72	.89	.418
Black	-8.23	.93	<.001	-8.41	1.20	<.001
Hispanic	-26.06	2.37	<.001	-31.16	3.84	<.001
Native language other than English	-28.49	4.10	<.001	-28.68	4.82	<.001
Time between start of school and testing	.14	.04	<.001	.20	.05	<.001
Has an IEP	-15.84	1.09	<.001	-16.76	1.52	<.001
TN-VPK participation	6.20	1.28	<.001	7.78	1.79	<.001
TN-VPK participation x Hispanic	21.88	3.91	<.001	24.42	5.93	<.001
TN-VPK participation x Non-English	29.48	5.20	<.001	34.04	7.05	<.001
Hispanic x Non-English	-16.51	5.09	.001	24.22	6.56	<.001
TN-VPK participation x Hispanic x Non-English	**-31.03**	6.97	**<.001**	**-42.45**	9.87	**<.001**

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom, and Hispanic children whose primary language was not English and who were in the treatment group served as the referent group for the parameter estimate. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table M-7
Statistical Models for the Effects of TN-VPK on the WJ Quantitative Concepts Subtest: Comparison of Children Who Were or Were Not Hispanic and Whose Primary Language was English or Another Language

Effect	WJ Quantitative Concepts 12 months			WJ Quantitative Concepts 6 months		
	b	SE	*p*	*b*	SE	*p*
Intercept	415.96	.78	<.001	415.96	1.04	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001
Region: West	-.90	.85	.294	-1.57	1.05	.136
Region: Central East	.20	.88	.823	.19	1.08	.862
Region: East	.55	.90	.543	1.53	1.09	.164
Male	-.92	.40	.020	-1.08	.58	.063
Black	-2.75	.58	<.001	-1.65	.80	.039
Hispanic	-8.09	1.49	<.001	-12.40	2.52	<.001
Native language other than English	-10.57	2.57	<.001	-11.35	3.16	<.001
Time between start of school and testing	.10	.02	<.001	.15	.03	<.001
Has an IEP	-7.92	.69	<.001	-7.74	3.88	.046
TN-VPK participation	7.27	.81	<.001	6.37	1.17	<.001
TN-VPK participation x Hispanic	5.02	2.46	.041	7.74	3.88	.046
TN-VPK participation x Non-English	16.28	3.27	<.001	20.16	4.61	<.001
Hispanic x Non-English	7.41	3.15	.019	12.17	4.29	.005
TN-VPK participation x Hispanic x Non-English	**-17.16**	4.37	**<.001**	**-24.80**	6.45	<.001

Note. Sample sizes are 4,144 for the 12-month bandwidth, 2,067 for the 6-month bandwidth, and 1,050 for the 3-month bandwidth. Only children eligible for free or reduced price lunch were included. Analyses were weighted to represent the statewide population of TN-VPK classrooms. Estimates were based on a multilevel model with children nested within pre-k classroom, and Hispanic children whose primary language was not English and who were in the treatment group served as the referent group for the parameter estimate. Previous analyses showed that the appropriate functional form for all models was linear.

Appendix Table N-1
Statistical Models for the Effects of TN-VPK on the WJ Composite:
Sample of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Composite 12 months			WJ Composite 6 months			WJ Composite 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	410.39	.82	<.001	411.75	1.09	<.001	413.03	1.46	<.001
Days from age eligibility cutoff	.04	<.00	<.001	.04	<.01	<.001	.04	.01	.008
Region: West	-1.37	.83	.102	-1.53	1.02	.135	-2.63	1.30	.044
Region: Central East	-.04	.84	.960	-.15	1.04	.889	-1.59	1.33	.234
Region: East	-.09	.86	.920	-.05	1.05	.962	-.37	1.35	.787
Male	-2.24	.35	<.001	-1.56	.51	.003	-2.22	.72	.003
Black	-3.35	.54	<.001	-2.57	.74	<.001	-1.73	.98	.076
Hispanic	-11.97	.91	<.001	-13.46	1.39	<.001	-14.22	1.98	<.001
Native language other than English	-6.51	1.02	<.001	-6.57	1.48	<.001	-7.65	2.09	<.001
Time between start of school and testing	.12	.02	<.001	.16	.03	<.001	.16	.04	<.001
Has an IEP	-10.08	.60	<.001	-10.27	.85	<.001	-10.69	1.20	<.001
Eligible for free or reduced price lunch	-4.50	.46	<.001	-5.98	.65	<.001	-6.42	.90	<.001
TN-VPK participation	**10.90**	.70	**<.001**	**10.82**	1.01	**<.001**	**10.30**	1.45	**<.001**

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. The composite is an average of the W-scores for the six tests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table N-2
Statistical Models for the Effects of TN-VPK on the WJ Letter-Word Subtest: Sample of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Letter-Word 12 months			WJ Letter-Word 6 months			WJ Letter-Word 3 months		
	B	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	336.30	1.44	<.001	338.69	1.87	<.001	340.76	2.58	<.001
Days from age eligibility cutoff	.04	<.01	<.001	.04	.01	<.001	.06	.02	.013
Region: West	-.67	1.46	.646	-1.58	1.77	.374	-2.56	2.36	.278
Region: Central East	-1.32	1.48	.372	-1.93	1.79	.283	-3.51	2.41	.147
Region: East	-1.79	1.51	.238	-1.71	1.82	.349	-2.13	2.44	.385
Male	-3.98	.61	<.001	-3.60	.86	<.001	-5.40	1.24	<.001
Black	1.02	.95	.283	3.27	1.26	.010	3.75	1.71	.029
Hispanic	-11.24	1.59	<.001	-12.16	2.37	<.001	-9.96	3.41	.004
Native language other than English	-2.29	1.77	.196	-4.14	2.52	.101	-5.00	3.60	.166
Time between start of school and testing	.23	.04	<.001	.29	.05	<.001	.32	.07	<.001
Has an IEP	-10.20	1.05	<.001	-11.28	1.45	<.001	-11.91	2.07	<.001
Eligible for free or reduced price lunch	-7.15	.81	<.001	-9.44	1.10	<.001	-9.50	1.55	<.001
TN-VPK Participation	**21.66**	1.22	**<.001**	**21.26**	1.71	**<.001**	**19.09**	2.50	**<.001**

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table N-3
Statistical Models for the Effects of TN-VPK on the WJ Spelling Subtest: Sample of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Spelling 12 months			WJ Spelling 6 months			WJ Spelling 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	368.55	1.26	<.001	370.56	1.77	<.001	371.74	2.39	<.001
Days from age eligibility cutoff	.05	<.01	<.001	.06	.01	<.001	.05	.02	.033
Region: West	1.09	1.22	.376	.55	1.66	.742	-1.06	2.12	.618
Region: Central East	.46	1.24	.711	-.06	1.68	.974	-2.29	2.18	.295
Region: East	1.88	1.27	.142	2.37	1.71	.167	1.81	2.21	.415
Male	-6.20	.57	<.001	-5.03	.82	<.001	-5.75	1.18	<.001
Black	-1932	.85	.024	-.85	1.20	.480	1.65	1.60	.303
Hispanic	-4.84	1.47	.001	-5.87	2.25	.009	-7.55	3.25	.020
Native language other than English	.23	1.64	.886	1.07	2.40	.655	1.04	3.44	.762
Time between start of school and testing	.21	.03	<.001	.24	.05	<.001	.27	.06	<.001
Has an IEP	-11.07	.97	<.001	-11.99	1.38	<.001	-11.81	1.98	<.001
Eligible for free or reduced lunch	-5.35	.74	<.001	-7.12	1.05	<.001	-8.28	1.48	<.001
TN-VPK Participation	**20.35**	1.14	**<.001**	**20.00**	1.63	**<.001**	**19.87**	2.39	**<.001**

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms. .

Appendix Table N-4
Statistical Models for the Effects of TN-VPK on the WJ Oral Comprehension Subtest: Sample of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Oral Comprehension 12 months			WJ Oral Comprehension 6 months			WJ Oral Comprehension 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	457.15	.85	<.001	458.22	1.11	<.001	459.27	1.39	<.001
Days from age eligibility cutoff	.03	<.01	<.001	.03	<.01	<.001	.02	.01	.103
Region: West	-3.47	.87	<.001	-2.92	1.06	.007	-3.52	1.21	.004
Region: Central East	.95	.88	.285	1.37	1.07	.203	.24	1.24	.848
Region: East	-.81	.90	.374	-.59	1.09	.588	-.73	1.26	.560
Male	-1.42	.35	<.001	-.71	.50	.159	-.63	.70	.372
Black	-5.45	.55	<.001	-5.35	.74	<.001	-5.68	.94	<.001
Hispanic	-11.54	.92	<.001	-14.08	1.38	<.001	-14.60	1.92	<.001
Native language other than English	-12.07	1.03	<.001	-11.30	1.47	<.001	-11.55	2.03	<.001
Time between start of school and testing	.05	.02	.018	.10	.03	<.001	.08	.04	.026
Has an IEP	-9.17	.60	<.001	-8.36	.84	<.001	-9.33	1.17	<.001
Eligible for free or reduced lunch	-2.98	.47	<.001	-4.49	.65	<.001	-4.75	.87	<.001
TN-VPK Participation	**3.86**	.71	**<.001**	**3.50**	1.00	**<.001**	**3.71**	1.42	**.009**

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table N-5
Statistical Models for the Effects of TN-VPK on the WJ Picture Vocabulary Subtest: Sample of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Picture Vocabulary 12 months			WJ Picture Vocabulary 6 months			WJ Picture Vocabulary 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	467.90	.78	<.001	468.58	1.04	<.001	470.51	1.37	<.001
Days from age eligibility cutoff	.02	<.01	<.001	0.02	<.01	<.001	.02	.01	.139
Region: West	-1.75	.78	.026	-1.32	.94	.162	-2.10	1.16	.074
Region: Central East	-.44	.79	.578	-0.34	.96	.724	-1.73	1.20	.151
Region: East	-.58	.81	.472	-1.03	.97	.289	-1.60	1.21	.190
Male	0.70	.35	.043	1.32	0.50	.009	1.10	.71	.119
Black	-2.11	.53	<.001	-2.42	0.71	<.001	-2.56	.93	.006
Hispanic	-18.58	.90	<.001	-20.34	1.37	<.001	-20.60	1.92	<.001
Native language other than English	-13.49	1.00	<.001	-12.83	1.46	<.001	-15.50	2.03	<.001
Time between start of school and testing	.03	.02	.187	0.03	0.03	.352	.01	.03	.725
Has an IEP	-7.36	.59	<.001	-6.98	0.84	<.001	-6.72	1.17	<.001
Eligible for free or reduced lunch	-2.40	.46	<.001	-3.06	.64	<.001	-3.73	.869	<.001
TN-VPK Participation	**3.40**	.70	<.001	**3.26**	1.00	**.002**	2.65	1.42	.063

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table N-6
Statistical Models for the Effects of TN-VPK on the WJ Applied Problems Subtest: Sample of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Applied Problems 12 months			WJ Applied Problems 6 months			WJ Applied Problems 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	412.37	1.27	<.001	413.53	1.58	<.001	415.18	2.13	<.001
Days from age eligibility cutoff	.05	<.01	<.001	0.04	0.01	<.001	0.04	0.02	.045
Region: West	-2.52	1.25	.045	-2.59	1.40	.068	-3.47	1.86	.064
Region: Central East	-.33	1.26	.793	0.36	1.43	.799	-.86	1.91	.654
Region: East	.27	1.30	.834	-0.52	1.45	.719	-1.19	1.93	.534
Male	-1857	.56	.001	-0.76	0.78	.327	-2.13	1.07	.047
Black	-8.59	.85	<.001	-8.49	1.09	<.001	-7.31	1.43	<.001
Hispanic	-18.67	1.46	<.001	-20.16	2.12	<.001	-24.36	2.93	<.001
Native language other than English	-9.50	1.63	<.001	-9.95	2.25	<.001	-11.20	3.09	<.001
Time between start of school and testing	.12	.03	<.001	0.16	0.04	<.001	0.11	0.05	.036
Has an IEP	-14.30	.96	<.001	-14.03	1.30	<.001	-14.88	1.78	<.001
Eligible for free or reduced lunch	-4.92	.74	<.001	-6.25	.98	<.001	-6.83	1.33	<.001
TN-VPK Participation	**8.53**	1.13	**<.001**	**9.93**	1.54	**<.001**	**9.52**	2.15	**<.001**

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table N-7
Statistical Models for the Effects of TN-VPK on the WJ Quantitative Concepts Subtest: Sample of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Quantitative Concepts 12 months			WJ Quantitative Concepts 6 months			WJ Quantitative Concepts 3 months		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	419.91	.83	<.001	420.93	1.10	<.001	421.03	1.47	<.001
Days from age eligibility cutoff	.03	<.01	<.001	0.04	<.01	<.001	.03	.01	.026
Region: West	-.79	.83	.344	-1.19	1.02	.247	-2.85	1.27	.027
Region: Central East	-.03	.84	.970	-.24	1.04	.820	-1.51	1.31	.251
Region: East	.61	.86	.476	.97	1.06	.364	1.30	1.33	.328
Male	-.74	.36	.040	-0.56	0.52	.278	-.49	.74	.511
Black	-3.06	.55	<.001	-1.92	0.75	.011	-.78	.99	.430
Hispanic	-7.15	.93	<.001	-8.59	1.42	<.001	-8.56	2.03	<.001
Native language other than English	-1.95	1.04	.061	-2.07	1.51	.170	-3.53	2.14	.100
Time between start of school and testing	.09	.02	<.001	0.13	0.03	<.001	.15	.04	<.001
Has an IEP	-8.33	.61	<.001	-8.78	.87	<.001	-9.24	1.23	<.001
Eligible for free or reduced price lunch	-4.06	.47	<.001	-5.41	.66	<.001	-5.45	.92	<.001
TN-VPK Participation	**7.55**	.72	**<.001**	**7.02**	1.03	<.001	**6.82**	1.49	**<.001**

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Estimates were based on a multilevel model with children nested within pre-k classroom. Previous analyses showed that the appropriate functional form for all models was linear. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table O-1
Statistical Models for the Effects of TN-VPK on the WJ Composite: Comparison of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	WJ Composite 12-month bandwidth			WJ Composite 6-month bandwidth			WJ Composite 3-month bandwidth		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	411.46	.89	<.001	412.21	1.20	<.001	412.72	1.61	<.001
TN-VPK participation (condition)	8.86	.97	<.001	9.98	1.38	<.001	10.91	1.97	<.001
Days from age eligibility cutoff	.04	.00	<.001	.04	.00	<.001	.04	.01	.008
Region: West	-1.31	.83	.115	-1.51	1.02	.141	-2.66	1.30	.042
Region: Central East	-.10	.84	.907	-.18	1.04	.862	-1.57	1.33	.239
Region: East	-.09	.86	.916	-.04	1.05	.971	-.38	1.35	.780
Male	-2.24	.35	<.001	-1.57	0.51	.002	-2.22	0.72	.002
Black	-3.38	.54	<.001	-2.59	.74	<.001	-1.71	.98	.080
Hispanic	-11.95	.91	<.001	-13.48	1.39	<.001	-14.20	1.98	<.001
Native language other than English	-6.54	1.02	<.001	-6.55	1.48	<.001	-7.68	2.09	<.001
Has an IEP	-10.09	.60	<.001	-10.29	.85	<.001	-10.70	1.20	<.001
Days between start of school and testing	.12	.02	<.001	.16	.03	<.001	0.16	.04	<.001
Free or reduced price lunch	-5.87	.65	<.001	-6.55	.90	<.001	-6.02	1.25	<.001
Condition x free or reduced price lunch	**2.65**	**.87**	**.002**	1.14	1.24	.358	-.80	1.74	.644

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. The composite is an average of the W-scores for the six tests used to measure outcomes: Letter-Word, Spelling, Oral Comprehension, Picture Vocabulary, Applied Problems, and Quantitative Concepts. All analyses were weighted.

Appendix Table O-2
Statistical Models for the Effects of TN-VPK on the WJ Letter-Word Subtest: Comparison of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	Letter-Word 12-month bandwidth			Letter-Word 6-month bandwidth			Letter-Word 3-month bandwidth		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	337.79	1.56	<.001	339.39	2.05	<.001	341.00	2.82	<.001
TN-VPK participation (condition)	18.82	1.69	<.001	19.92	2.35	<.001	18.61	3.38	<.001
Days from age eligibility cutoff	.04	.00	<.001	.04	.01	<.001	.06	.02	.013
Region: West	-.60	1.45	.679	-1.54	1.76	.385	-2.54	2.36	.283
Region: Central East	-1.39	1.47	.345	-1.98	1.79	.270	-3.53	2.41	.145
Region: East	-1.79	1.51	.237	-1.69	1.82	.354	-2.12	2.45	.387
Male	-3.98	.61	<.001	-3.62	.86	<.001	-5.40	1.24	<.001
Black	.99	.95	.293	3.25	1.26	.010	3.73	1.71	.030
Hispanic	-11.21	1.59	<.001	-12.19	2.37	<.001	-9.98	3.41	.004
Native language other than English	-2.33	1.77	.190	-4.11	2.52	.103	-4.97	3.61	.169
Has an IEP	-10.22	1.04	<.001	-11.31	1.45	<.001	-11.90	2.07	<.001
Days between start of school and testing	.22	.04	<.001	.29	.05	<.001	.32	.07	<.001
Free or reduced price lunch	-9.06	1.12	<.001	-10.33	1.54	<.001	-9.81	2.15	<.001
Condition x free or reduced price lunch	**3.70**	**1.51**	**.015**	1.75	2.10	.406	.63	2.99	.834

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table O-3
Statistical Models for the Effects of TN-VPK on the WJ Spelling Subtest: Comparison of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	Spelling 12-month bandwidth			Spelling 6-month bandwidth			Spelling 3-month bandwidth		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Intercept	369.00	1.38	<.001	370.86	1.94	<.001	370.46	2.64	<.001
TN-VPK participation (condition)	19.50	1.56	<.001	19.43	2.24	<.001	22.36	3.24	<.001
Days from age eligibility cutoff	.05	.00	<.001	.06	.01	<.001	.05	.02	.033
Region: West	1.11	1.22	.366	.56	1.66	.734	-1.17	2.12	.583
Region: Central East	.44	1.24	.726	-.08	1.68	.963	-2.21	2.17	.310
Region: East	1.88	1.27	.142	2.38	1.71	.167	1.77	2.20	.424
Male	-6.20	.57	<.001	-5.04	.82	<.001	-5.74	1.18	<.001
Black	-1.94	.85	.023	-.86	1.20	.474	1.75	1.60	.278
Hispanic	-4.83	1.47	.001	-5.89	2.25	.009	-7.48	3.25	.022
Native language other than English	.22	1.64	.892	1.08	2.40	.652	.90	3.44	.794
Has an IEP	-11.08	.97	<.001	-12.00	1.38	<.001	-11.86	1.98	<.001
Days between start of school and testing	.21	.03	<.001	.24	.05	<.001	.27	.06	<.001
Free or reduced price lunch	-5.92	1.04	<.001	-7.50	1.46	<.001	-6.64	2.06	.001
Condition x free or reduced price lunch	1.11	1.40	.428	.75	2.00	.708	-3.26	2.85	.253

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table O-4
Statistical Models for the Effects of TN-VPK on the WJ Oral Comprehension Subtest: Comparison of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	Oral Comprehension 12-month bandwidth			Oral Comprehension 6-month bandwidth			Oral Comprehension 3-month bandwidth		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	457.20	.92	<.001	457.89	1.21	<.001	458.17	1.54	<.001
TN-VPK participation (condition)	3.77	.97	<.001	4.12	1.37	.003	5.90	1.92	.002
Days from age eligibility cutoff	.03	.00	<.001	.03	.00	<.001	.02	.01	.102
Region: West	-3.47	.87	<.001	-2.94	1.06	.006	-3.61	1.20	.003
Region: Central East	.94	.88	.286	1.40	1.07	.195	.30	1.23	.808
Region: East	-.81	.90	.374	-.60	1.09	.583	-.78	1.25	.531
Male	-1.42	.35	<.001	-.70	.50	.162	-.62	.70	.375
Black	-5.45	.55	<.001	-5.34	.74	<.001	-5.61	.94	<.001
Hispanic	-11.54	.92	<.001	-14.06	1.38	<.001	-14.52	1.92	<.001
Native language other than English	-12.07	1.03	<.001	-11.31	1.47	<.001	-11.68	2.03	<.001
Has an IEP	-9.17	.60	<.001	-8.35	.85	<.001	-9.36	1.17	<.001
Days between start of school and testing	.05	.02	.018	.10	.03	<.001	.08	.04	.021
Free or reduced price lunch	-3.04	.65	<.001	-4.07	.90	<.001	-3.32	1.22	.007
Condition x free or reduced price lunch	.12	.87	.887	-.82	1.23	.505	-2.85	1.69	.092

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table O-5
Statistical Models for the Effects of TN-VPK on the WJ Picture Vocabulary Subtest: Comparison of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	Picture Vocabulary 12-month bandwidth			Picture Vocabulary 6-month bandwidth			Picture Vocabulary 3-month bandwidth		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	468.90	.86	<.001	469.18	1.16	<.001	471.04	1.53	<.001
TN-VPK participation (condition)	1.51	.96	.113	2.13	1.37	.120	1.60	1.93	.407
Days from age eligibility cutoff	.02	.00	<.001	.02	.00	<.001	.02	.01	.140
Region: West	-1.71	.78	.032	-1.29	.95	.173	-2.06	1.17	.081
Region: Central East	-.49	.79	.532	-.38	.96	.689	-1.76	1.20	.146
Region: East	-.59	.81	.470	-1.01	.98	.300	-1.57	1.22	.201
Male	.70	.35	.042	1.31	.50	.009	1.10	.71	.121
Black	-2.13	.53	<.001	-2.44	.71	<.001	-2.59	.93	.006
Hispanic	-18.56	.90	<.001	-20.37	1.37	<.001	-20.63	1.93	<.001
Native language other than English	-13.51	1.00	<.001	-12.81	1.46	<.001	-15.44	2.04	<.001
Has an IEP	-7.37	.59	<.001	-7.00	.84	<.001	-6.71	1.17	<.001
Days between start of school and testing	.03	.02	.209	.02	.03	.357	.01	.03	.756
Free or reduced price lunch	-3.68	.64	<.001	-3.81	.89	.927	-4 .42	1.22	<.001
Condition x free or reduced price lunch	**2.47**	**.86**	**.004**	1.48	1.22	.225	1.37	1.70	.421

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table O-6
Statistical Models for the Effects of TN-VPK on the WJ Applied Problems Subtest: Comparison of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	Applied Problems 12-month bandwidth			Applied Problems 6-month bandwidth			Applied Problems 3-month bandwidth		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	414.87	1.38	<.001	414.48	1.76	<.001	414.86	2.35	<.001
TN-VPK participation	3.80	1.55	.014	8.14	2.11	<.001	10.16	2.92	<.001
Days from age eligibility cutoff	.05	.00	<.001	.04	.00	<.001	.04	.02	.045
Region: West	-2.39	1.24	.056	-2.53	1.40	.073	-3.49	1.86	.062
Region: Central East	.20	1.26	.875	.29	1.43	.839	-.84	1.91	.661
Region: East	.26	1.29	.841	-.50	1.45	.730	-1.20	1.93	.535
Male	-1.85	.56	.001	-.77	.78	.319	-2.13	1.07	.047
Black	-8.66	.85	<.001	-8.54	1.09	<.001	-7.29	1.43	<.001
Hispanic	-18.62	1.46	<.001	-20.22	2.12	<.001	-24.34	2.93	<.001
Native language other than English	-.9.56	1.62	<.001	-9.91	2.25	<.001	-11.23	3.10	<.001
Has an IEP	-14.34	.96	<.001	-14.06	1.30	<.001	-14.89	1.78	<.001
Days between start of school and testing	.11	.03	<.001	.16	.04	<.001	.12	.05	.035
Free or reduced price lunch	-8.12	1.03	<.001	-7.45	1.38	<.001	-6.42	1.85	<.001
Condition x free or reduced price lunch	**6.17**	**1.39**	**<.001**	2.34	1.89	.215	-.83	2.58	.747

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table O-7
Statistical Models for the Effects of TN-VPK on the WJ Quantitative Concepts Subtest: Comparison of Children Who Were or Were Not Eligible for Free or Reduced Price Lunch

Effect	Quantitative Concepts 12-month bandwidth			Quantitative Concepts 6-month bandwidth			Quantitative Concepts 3-month bandwidth		
	b	**SE**	***p***	***b***	**SE**	***p***	***b***	**SE**	***p***
Intercept	420.86	.90	<.001	421.46	1.22	<.001	420.99	1.62	<.001
TN-VPK participation	5.73	.99	<.001	6.01	1.41	<.001	6.90	2.03	<.001
Days from age eligibility cutoff	.03	.00	<.001	.04	.00	<.001	.03	.01	.026
Region: West	-.74	.82	.372	-1.16	1.02	.259	-2.85	1.27	.027
Region: Central East	-.08	.83	.925	-.28	1.04	.790	-1.51	1.31	.252
Region: East	.61	.86	.474	.97	1.05	.357	1.30	1.33	.329
Male	-.74	.36	.040	-.57	.52	.271	-.49	.74	.511
Black	-3.08	.55	<.001	-1.94	.75	.010	-.78	.99	.433
Hispanic	-7.14	.93	<.001	-8.62	1.42	<.001	-8.56	2.03	<.001
Native language other than English	-1.97	1.04	.058	-2.05	1.51	.174	-3.53	2.14	.010
Has an IEP	-8.34	.61	<.001	-8.80	.87	<.001	-9.24	1.23	<.001
Days between start of school and testing	.09	.02	<.001	.13	.03	<.001	.15	.04	<.001
Free or reduced price lunch	-5.29	.66	<.001	-6.08	.92	<.001	-5.40	1.28	<.001
Condition x free or reduced price lunch	**2.38**	**.89**	**.007**	1.32	1.26	.296	.10	1.78	.956

Note. Sample sizes are 5,188 for the 12-month bandwidth, 2,602 for the 6-month bandwidth, and 1,334 for the 3-month bandwidth. Analyses were weighted to represent the statewide population of TN-VPK classrooms.

Appendix Table P-1

Effect Size Estimates from Age-Cutoff Regression-Discontinuity Studies of State and Locally-Funded Pre-K Programs

State or Locality	Study	Measure	Outcome			
			Literacy	Language	Math	Other
Arkansas	Jung, Barnett, Hustedt, & Francis (2013)	CTOPPP Print Awareness	.82*			
		PPVT		.28*		
		WJ Applied Problems			.33*	
Boston, Massachusetts	Weiland & Yoshikawa (2013)	WJ Letter-Word Identification	.62*			
		PPVT		.44*		
		WJ Applied Problems			.59*	
		Early Math Assessment			.43*	
		Pencil Tap				.21*
		Backward Digit Span				.24*
		Forward Digit Span				.24*
		Dimensional Card Sort				.28*
		TOQ Attention				.11*
California	Barnett, Howes, & Jung (2009)	CTOPPP Print Awareness	1.17*			
		PPVT		.30*		
		WJ Applied Problems			.38*	
Connecticut	Connecticut Academy of Science and Engineering (2016)	WJ Letter-Word Identification & Word Attack (composite)	.69*			
		PPVT		.05		
		WJ Oral Comprehension		.31		
		WJ Applied Problems, Calculations, & Math Fluency (composite)			.48*	
		Behavior Assessment Scale for Children (BASC-3)				n.s
Georgia	Peisner-Feinberg, Schaaf, LaForett, Hildebrandt, & Sideris (2014)	Letter Knowledge	.89*			
		WJ Letter-Word Identification	1.05*			
		WJ Sound Awareness	.59*			
		WJ Word Attack	1.20*			
		WJ Picture Vocabulary		.01		
		WJ Applied Problems			.51*	
		Counting			.86*	
		Social Awareness				.43*
		SSiS Social Skills				.23
		SSiS Problem Behavior				.10

continued

Appendix Table P-1 (continued)
Effect Size Estimates from Age-Cutoff Regression-Discontinuity Studies of State and Locally-Funded Pre-K Programs

State or Locality	Study	Measure	Outcome			
			Literacy	Language	Math	Other
Kalamazoo County Michigan	Bartik (2013)	WJ Letter-Word Identification	.25			
		PPVT		.60*		
		WJ Applied Problems			.70*	
		Devereux Early Childhood Assessment Protective Factors				.51
		Devereux Early Childhood Assessment Behavioral Concerns				-.28
Los Angeles, California	Kyger & Barnhart (2017)	DIBELS Letter Naming Fluency	.83*			
		DIBELS Phoneme Segmentation Fluency		.76*		
		DIBELS Initial Sound Fluency		.43		
		IGDis-EN Oral Counting			.08	
		IGDis-EN Quantity Comparison			-.50	
		IGDis-EN Correspondence Counting			1.83	
Massachusetts	Hofer, Checkoway, Goodson, & Nichols (2018)	WJ Letter-Word Identification	.92*			
		PPVT		.21*		
		WJ Applied Problems			.45*	
		Executive Functioning (hearts & flowers task)				.05
Michigan	Wong, Cook, Barnett, & Jung (2008)	CTOPPP Print Awareness	1.09*			
		PPVT		-.13		
		WJ Applied Problems			.53*	
New Jersey	Wong, Cook, Barnett, & Jung (2008)	CTOPPP Print Awareness	.32*			
		PPVT		.36*		
		WJ Applied Problems			.23*	
New Mexico	Hustedt, Barnett, Jung, & Friedman (2010)	ELSA Early Literacy	1.30*			
		PPVT		.24*		
		WJ Applied Problems			.37*	
North Carolina	Peisner-Feinberg & Schaaf (2011)	Phonological Awareness (TOPEL)	.56*			
		Print Knowledge (TOPEL)	1.16*			
		WJ Letter-Word Identification	1.14*			
		PPVT		.06		
		WJ Applied Problems			.34*	
		Counting			.81*	
Oklahoma	Wong, Cook, Barnett, & Jung (2008)	CTOPPP Print Awareness	.42			
		PPVT		.28*		
		WJ Applied Problems			.34	

continued

Appendix Table P-1 (*continued)*
Effect Size Estimates from Age-Cutoff Regression-Discontinuity Studies of State and Locally-Funded Pre-K Programs

State or City	Study	Measure	Outcome			
			Literacy	Language	Math	Other
San Francisco	Applied Survey Research (2013)	WJ Letter-Word Identification	.40*			
		Receptive One-Word Picture Vocabulary (ROWPVT)		.29		
		WJ Applied Problems			.40*	
		Self-Regulation (HTKS)				.51*
South Carolina	Wong, Cook, Barnett, & Jung (2008)	CTOPPP Print Awareness	.78*			
		PPVT		.04		
Tennessee	Pion & Lipsey (2020)	WJ Composite				.83*
		WJ Letter-Word Identification	.99*			
		WJ Spelling	.97*			
		WJ Oral Comprehension		.28*		
		WJ Picture Vocabulary		.32*		
		WJ Applied Problems			.47*	
		WJ Quantitative Concepts			.52*	
		Note: Estimates for full sample				
Tennessee	Coburn (2009)	Brigance Preschool & K-1 Screen				NA
Tulsa, Oklahoma	Gormley, Gayer, Philips, & Dawson (2005) Gormley & Gayer (2005) Gormley, Phillips, & Gayer (2008) Phillips & Meloy (2012) Gormley (2008)	WJ Letter-Word Identification	.79*			
		WJ Spelling	.64*			
		WJ Applied Problems			.38*	
		Cognitive/knowledge				.39*
		Language		.38*		
		Motor Skills				.24*
		Social-emotional				n.s.
Virginia	Huang (2017)	PALS Lowercase Letter Recognition	.95*			
		Note: Midpoint of reported range				
Virginia	Gaylor, Golan, Chow, Grindal, Mercier, Williamson, & Tiruke (2019)	PALS Letter Sounds	1.04*			
		PALS Lowercase Letter Recognition	1.04*			
		PPVT		.23*		
		WJ Applied Problems			.35*	
		Self-regulation (HTKS)				.31*
		Note: Mean of 2 estimates				
West Virginia	Wong, Cook, Barnett, & Jung (2008)	CTOPPP Print Awareness	.92*			
		PPVT		.16		
		WJ Applied Problems			.06	

Notes. CTOPPP= Comprehensive Test of Phonological and Print Processing. PALS=Phonological Awareness Literacy Screening. PPVT=Peabody Picture Vocabulary Test. SSiS=Social Skills Improvement System. TOPEL=Test of Preschool Early Literacy. TOQ=Task Orientation Questionnaire. WJ=Woodcock Johnson III Achievement Battery.
* $p<.05$

Appendix P-2: References for Studies in Table P-1

Arkansas

Jung, K., Barnett, W. S., Hustedt, J. T., & Francis, J. (2013). *Longitudinal effects of the Arkansas Better Chance Program: Findings from first grade through fourth grade.* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Hustedt, J.T., Barnett, W.S., Jung, K., & Thomas, J. (2007). The effects of the Arkansas Better Chance Program on young children's school readiness. (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Boston, Massachusetts

Weiland, C. & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84(6),* 2112–2130.

Weiland, C. (undated). *Impacts of the Boston prekindergarten program on the school readiness of young children with special needs*. Manuscript.

California

Barnett, W. S., Howes, C., & Jung, K. (2009). *California's State Preschool Program: Quality and Effects on Children's Cognitive Abilities at Kindergarten Entry* (Report to the California Children and Families Commission). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Connecticut

Connecticut Academy of Science and Engineering. (2016). *Early Childhood Regression Discontinuity Study*.(Report for the Connecticut General Assembly and the Education Committee). Rocky Hill, CT: Connecticut Academy of Science and Engineering.

Georgia

Peisner-Feinberg, E. S., Schaaf, J. M., LaForett, D. R., Hildebrandt, L.M., & Sideris, J. (2014). *Effects of Georgia's Pre-K Program on Children's School Readiness Skills: Findings from the 2012-2013 Evaluation Study*. Chapel Hill: The University of North Carolina, FPG Child Development Institute.

Kalamazoo County, Michigan

Bartik, T. J. (2013). *Effects of the Pre-K Program of Kalamazoo County Ready 4s on Kindergarten Entry Test Scores: Estimates Based on Data from the Fall of 2011 and the Fall of 2012* (Upjohn Institute Working Paper 13-198). Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Los Angeles, California

Kyger, M., & Barnhart, M. (2017). *Effects of High-Quality Preschool on Early Literacy and Mathematics Skills—A Regression-Discontinuity Approach.* Los Angeles, CA: Los Angeles Universal Preschool, Research and Evaluation.

Massachusetts

Hofer, K. G., Checkoway, A., Goodson, B., & Nichols, A. (2018). *Massachusetts Preschool Expansion Grant (PEG) Impact Evaluation Report* (Research Report). Cambridge, MA: Abt Associates, Inc.

Michigan

Lamy, C., Barnett, W. S., & Jung, K. (2005a). *The Effects of the Michigan School Readiness Program on Young Children's Abilities at Kindergarten Entry* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, *27*, 122–154. doi:10.1002/pam.2031

New Jersey

Lamy, C., Barnett, W. S., & Jung, K. (2005). *The Effects of New Jersey's Abbott Preschool Program on Young Children's School Readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, *27*, 122–154. doi:10.1002/pam.2031

New Mexico

Hustedt, J. T., Barnett, W. S., Jung, K., & Friedman, A. H. (2010). *The New Mexico Pre-k Evaluation: Impacts from the Fourth Year (2008-2009) of New Mexico's State-funded Pre-k Program* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

North Carolina

Peisner-Feinberg, E. S., & Schaaf, J. M. (2011). *Summary of Key Findings: Effects of the North Carolina More at Four Prekindergarten Program on Children's School Readiness Skills*. Chapel Hill: The University of North Carolina, FPG Child Development Institute.

Oklahoma

Lamy, C., Barnett, W. S., & Jung, K. (2005). *The Effects of Oklahoma's Early Childhood Four-year-old Program on Young Children's School Readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, *27*, 122–154. doi:10.1002/pam.2031

San Francisco

Applied Survey Research (2013, August). *Evaluating Preschool for All Effectiveness*. Research Brief. San Jose, CA: Applied Survey Research.

South Carolina

Lamy, C., Barnett, W. S., & Jung, K. (2005). *The effects of South Carolina's Early Childhood Programs on Young Children's School Readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, *27*, 122–154. doi:10.1002/pam.2031

Tennessee

Pion, G. M., & Lipsey, M. W. (2020). *The Estimated Impact of the Tennessee Voluntary Prekindergarten Program on Children's Literacy, Language, and Mathematics Skills: Results from a Regression-Discontinuity Design.* Nashville, TN: Peabody Research Institute, Vanderbilt University. [This report]

Coburn, J.L. (2009). *The Effects of Tennessee's Prekindergarten Programs on Young Children's School Readiness Skills: A Regression-Discontinuity Design*. Dissertation. Tennessee Technological University.

Tulsa, Oklahoma

Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-k on cognitive development. *Developmental Psychology, 41(6),* 872-884.

Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. *The Journal of Human Resources, 40(3),* 533-558.

Gormley, W. T., Phillips, D., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science, 320(June 27),* 1723-1724.

Phillips, D.A., & Meloy, M.C. (2012). High-quality school-based pre-k can boost early learning for children with special needs. *Exceptional Children, 78(4),* 471-490.

Gormley, W.T. (2008). The effects of Oklahoma's pre-k program on Hispanic children. *Social Science Quarterly, 89(4),* 916-936.

Virginia

Huang, F. L. (2017). Does attending a state-funded preschool program improve letter name knowledge? *Early Childhood Research Quarterly, (38),* 116-126.

Gaylor, E., Golan, S., Chow, K., Grindal, T., Mercier, B., Williamson, C., & Tiruke, T. (2019). *Virginia Preschool Initiative Plus (VPI+): Final evaluation report*. Prepared for The Virginia Department of Education. Menlo Park, CA: SRI International.

West Virginia

Lamy, C., Barnett, W. S., & Jung, K. (2005). *The Effects of West Virginia's Early Education Program on Young Children's School Readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, *27*, 122–154. doi:10.1002/pam.2031