Washington State Institute for Public Policy
110 Fifth Avenue SE, Suite 214 • PO Box 40999 • Olympia, WA 98504 • 360.664.9800 • www.wsipp.wa.gov

January 2022

Evaluation of the Early Childhood Education and Assistance Program: *Short- and Long-Term Outcomes for Children*

The Early Childhood Education and Assistance Program (ECEAP) provides preschool education and comprehensive services to low-income children and their families in Washington State.

The 2019 Washington State Legislature directed WSIPP to explore short and long-term outcomes related to ECEAP enrollment. This legislation also directed us to examine "the effects of full-day programming and half-day programming on outcomes."[1]

This report focuses on the first part of this assignment, detailing the short and long-term outcomes for program participants, including high school graduation.

Section I provides an overview of WSIPP's portfolio of early education research. Section II provides background on ECEAP. Section III details our data and sample construction. Section IV outlines our evaluation methodology. Section V describes our findings. Section VI summarizes the results and limitations of our evaluation.

Summary

The 2019 Washington State Legislature directed WSIPP to evaluate the impact of Early Childhood Education and Assistance Program (ECEAP), including long-term outcomes such as high school graduation. This follow-up includes long-term outcome results for the previously studied historical cohort group (children born between 1996-2004), as well as more short-term outcomes for the historical cohort and a more recent cohort group (children born between 2004-2014).

We found the relationship between ECEAP participation and outcomes is strongest immediately after ECEAP participation and (generally) fades over time.

Compared to eligible children who did not participate in ECEAP, ECEAP participants are—

- More likely to be kindergarten-ready and
- Less likely to participate in special education in early school years.

However, we did not find clear evidence that ECEAP participants had better or worse outcomes compared to non-participants on the following:

- 3rd to 5th grade assessments,
- Criminal convictions during high school, and
- High school graduation.

These findings are consistent with recent literature on the impact of early childhood education programs but differ from our findings from an earlier evaluation of ECEAP. We discuss changes in our methodology in and implications of these findings in Appendix VI.

[1] Engrossed Second Substitute House Bill 1391, Chapter 369, Laws of 2019. Section 7(5).

I. WSIPP's Early Childhood Education Research Portfolio

High-quality early childhood education (ECE) programs have been shown to positively impact child developmental, academic, and behavioral outcomes as well as parental outcomes including maternal employment and earnings.[2] However, ECE programs vary, and results are often mixed depending on program models and populations served. Further research evidence indicates that effects might be limited and that early effects often fade as children progress through school.

As knowledge about the potential benefits of ECE programs has grown, so too has interest and public investment in issues like affordable childcare, universal preschool,[3] supports for the ECE workforce, and rating systems that establish quality standards for childcare and educational providers.[4] Investments in ECE in Washington State have increased over time too, resulting in WSIPP's current portfolio of ECE research projects.

The Washington State Legislature has directed WSIPP to evaluate several early childhood education programs in Washington State. This includes evaluations of the Early Childhood Education and Assistance Program (ECEAP), a preschool program for low-income children, and Early Achievers, the state's quality rating and improvement system that many childcare and education programs are required to participate in. We offer a summary of the ECEAP evaluations to provide context for this report and its relevance in WSIPP's collection of ECE research.[5]

WSIPP's Review of Evidence on ECE Programs

In 2014, WSIPP released a meta-analysis of research on Early Childhood Education (ECE) for low-income children. As part of the report, WSIPP conducted a benefit-cost analysis to determine if the measured benefits of the program exceeded program costs. Our report found evidence that ECE programs can improve outcomes and that the benefits of such programs were likely to exceed costs.[6]

In 2019, WSIPP updated this meta and cost-benefit analyses. The report narrowed its review to only include studies evaluating programs after 1975, due to meaningful changes to societal context, program models, and evaluation methods after early evaluations were conducted. We found evidence that ECE programs improved some academic outcomes and that benefits were likely to exceed costs.[7]

[2] Hoagland, C., Fumia, D., & Reynolds, M. (2019). *Early childhood education for low-income students: A review of the evidence and benefit-cost analysis UPDATE* (Doc. No. 19-12-2201). Olympia: Washington State Institute for Public Policy.
[3] Refers to programs that provide preschool access to all children in a district or state.
[4] Friedman-Krausee, A.H., Barnett, W.S., Godges, K.S., Weisenfeld, G.G., & Gardiner, B.A. (2021). *The state of preschool 2020.* The National Institute for Early Education Research: New Jersey.
[5] WSIPP is also completing a series of evaluations of Early Achievers, the state's quality rating and improvement system (QRIS). Goodvin, R., Rashid, A., & He, L. (2020). *Early Achievers evaluation report two: Prekindergarten quality and child outcomes in kindergarten* (Doc. No. 20-12-2203). Olympia: Washington State Institute for Public Policy.
[6] For full results, see Hoagland et al. (2019).
[7] Ibid.

WSIPPs' Evaluations of ECEAP

In 2014, WSIPP released a report examining the outcomes of children who were enrolled in ECEAP in the early 2000s.[8] This report showed a positive relationship between ECEAP enrollment and children's 3rd-, 4th-, and 5th-grade test scores.

In 2019, again the Washington State Legislature directed WSIPP to evaluate the ECEAP program.

The first part of the legislative assignment was to update its previous evaluation and examine short-term and long-term outcomes, which is the focus of this report.

In this evaluation, we examine the following research questions relative to eligible children who were not enrolled in ECEAP:

I. What were the short-term and long-term academic outcomes for ECEAP participants?
II. What were the long-term behavioral outcomes for ECEAP participants?
III. Is ECEAP participation associated with changes in parental employment outcomes while the child is enrolled in ECEAP?

The second part of the legislative assignment, which directed WSIPP to evaluate "the effects of full-day programming and half-day programming," is addressed in a separate report. In that report, we estimate the relationship between enrollment in a School-Day program, relative to a Part-Day program, and child outcomes including kindergarten readiness. We find that there is a positive relationship between School-Day enrollment and children's kindergarten readiness. On average, children enrolled in School-Day classes are six percentage points more likely to demonstrate kindergarten readiness, compared to children in Part Day.[9]

See Exhibit 1 for WSIPP's legislative assignment and Exhibit 2 for a summary of how these reports fit together.

Exhibit 1
Legislative Assignment

> *. . . the Washington state institute for public policy shall update the outcome evaluation of the early childhood education and assistance program required by chapter 16, Laws of 2013 and report to the governor and the legislature on the outcomes of program participants. The evaluation must include the demographics of program participants including race, ethnicity, and socioeconomic status.* ***The evaluation must examine short and long-term impacts of program participants, including high school graduation rates for up to two cohorts.*** *When conducting the evaluation, the institute must consider, to the extent that data is available, the education levels and demographics, including race, ethnicity, and socioeconomic status of early childhood education and assistance program staff and the effects of full-day programming and half-day programming on outcomes.*
>
> Engrossed Second Substitute House Bill 1391, Chapter 369, Laws of 2019.

[8] Bania, N., Kay, N., Aos, S., & Pennucci, A. (2014). *Outcome evaluation of Washington State's Early Childhood Education and Assistance Program*. (Document No. 14-12-2201). Olympia: Washington State Institute for Public Policy.

[9] For full results see Cramer, J., & Rashid, A. (2022). *Evaluation of the Early Childhood Education and Assistance Program: Kindergarten readiness for school-and part-day enrollees.* (Doc. No. 22-01-2201). Olympia: Washington State Institute for Public Policy.

Exhibit 2

Early Childhood Education Report Series

	Report one Meta-analysis of ECE programs	Report two Outcome evaluation of ECEAP	Report three Updated meta-analysis of ECE programs	Report four Outcome evaluation of ECEAP	Report five Outcome evaluation of ECEAP (dosage impacts)
Report overview	Systematic review of research on academic, social, and emotional development outcomes for children in ECE programs.	Retrospective evaluation of ECEAP.	Update of the 2014 systematic review focusing on research on academic, social, and emotional development outcomes for children in ECE programs.	Update to the 2014 outcome evaluation and focused on short- and long-term outcomes.	Evaluation of ECEAP dosage models (Part Day vs School Day).
Population evaluated	Children eligible to participate in ECE programs in the U.S. Treatment groups include children in ECE programs. Control groups include non-ECE participants (some studies assessed treatment as usual so control groups may have included children in other ECE programs).	Children born September 1999 – August 2004 who received Basic Food benefits when they were three or four years old. Children in ECEAP were the treatment group. Children who were not in ECEAP were the comparison group.	Children eligible to participate in ECE programs in the U.S. Treatment groups include children in ECE programs. Control groups include non-ECE participants (some studies assessed treatment as usual so control groups may have included children in other ECE programs).	Children who received DSHS services when they were three or four years old. Historical cohort: children born between September 1996 – August 2004. Recent cohort: children born between September 2004 – August 2014. Children who received ECEAP were in the treatment group. Children who were similar but did not receive ECEAP were in the comparison group.	Children enrolled in ECEAP between academic years 2014-15 and 2018-19. The treatment group is comprised of children enrolled in School-Day classes. The comparison group is children enrolled in Part-Day classes.
Outcomes	Test scores, high school graduation, grade retention, special education placement, criminal justice involvement, teen births, and self-regulation.	Reading and math test scores measured in 3rd, 4th, and 5th grades.	Test scores, grade retention, special education placement, attendance, GPA, high school graduation, and college enrollment.	Kindergarten readiness, test scores, special education placement, high school graduation, criminal justice involvement, parental employment, Child Protective Services (CPS) involvement.	Kindergarten readiness, special education placement, monthly absences.
Published	January 2014	December 2014	December 2019	January 2022	January 2022

II. Background

The Early Childhood Education and Assistance Program (ECEAP) was created in 1985 and is Washington's preschool program. Modeled after Head Start, ECEAP is designed to support children ages three and four who are eligible based on family income, developmental need, and/or other factors.

Administered by the Department of Children, Youth, and Families (DCYF), ECEAP focuses on early education to support children's social-emotional and pre-academic development. Providers also offer wraparound health and nutrition services for children and family engagement.[10] For example, children are provided a traditional classroom preschool education and receive developmental screenings, periodic assessments, and individualized support to prepare them for kindergarten. Children also receive daily nutritious meals, medical and dental screenings, and mental health care referrals as needed.

In terms of family engagement, ECEAP staff connect with parents and guardians to provide resources and support financial and housing stability, employment, and educational attainment.

As of 2020, over 390 ECEAP providers were operating in various settings including public schools, childcare centers and homes, tribal organizations, community colleges, and non-profits. Statewide, these sites served children in approximately 14,000 slots.[11]

Eligibility and Enrollment

Most children are eligible for ECEAP services for the following reasons:[12]

- They are at least three years old, but less than five years old, by August 31st of the school year they enroll, and
- Their family income is less than or equal to 110% of the federal poverty level (FPL).[13]

Children are also eligible if they qualify for special education services. Additionally, a limited number of children who live in families with incomes greater than 110% FPL and have certain research-based prioritization factors are also eligible.[14]

[10] Washington State Department of Children, Youth, & Families website *Early Childhood Education and Assistance Program.*
[11] Ibid.
[12] 2020-21 ECEAP Performance Standards.
[13] The equivalent of an annual income of $29,150 for a family of four in 2021. *2021 ECEAP Income Eligibility Limits.*
[14] These factors include environmental circumstances such as family violence, chemical dependency, child protective service involvement, incarcerated parents, foster care placement, and/or homelessness. 2020-21 ECEAP Performance Standards.

Changes to ECEAP

ECEAP has implemented several changes since its first inception. It has expanded the number of statewide available slots from 1,000 in 1985 to approximately 14,000 in 2020.[15] It has also increased the number of hours of instruction available to enrollees.[16]

It has also made several programmatic changes including requiring sites to participate in Early Achievers, a statewide quality rating and improvement system.

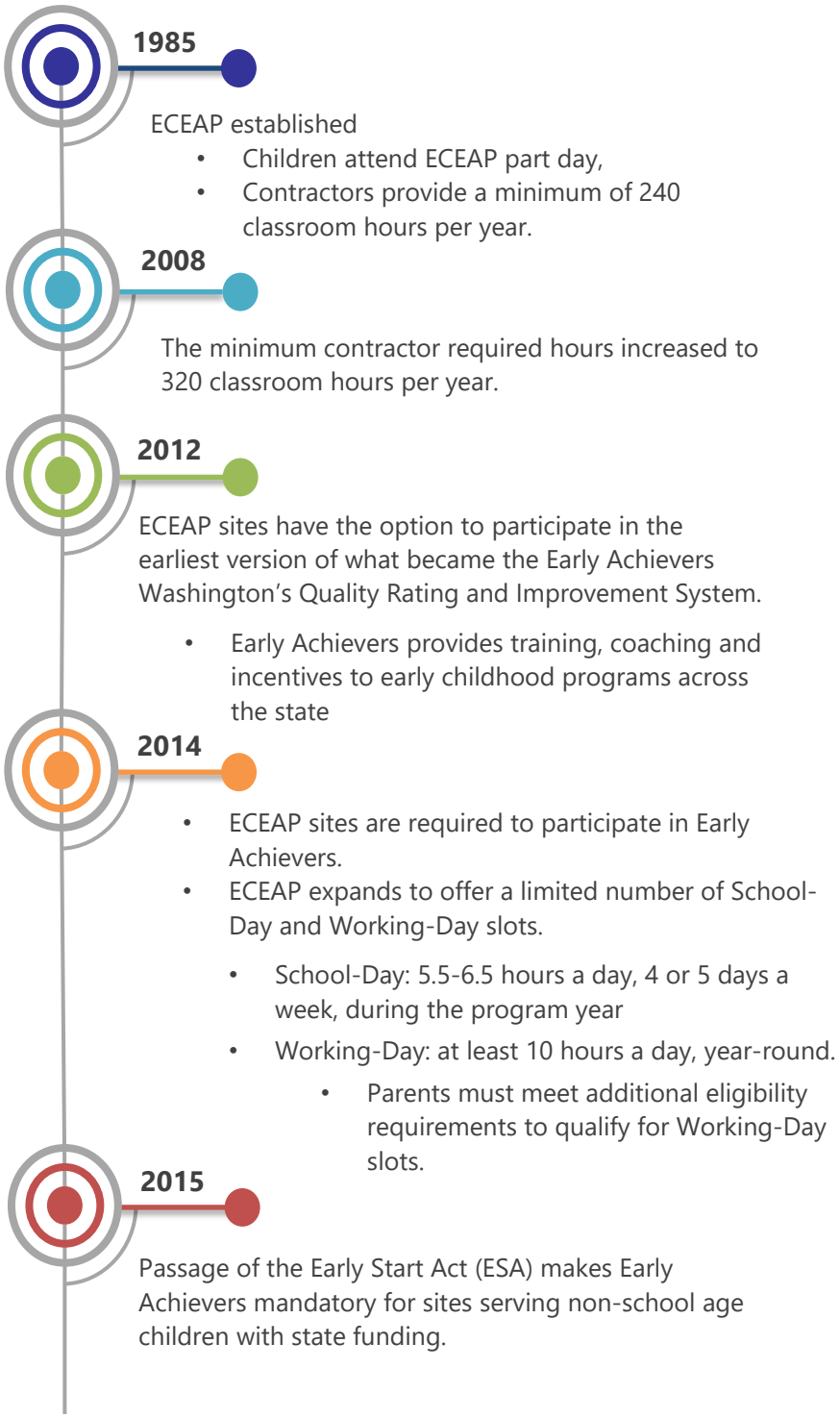
Exhibit 3 highlights key changes to the ECEAP program by academic year. The specific impact of the introduction of the QRIS system in 2012, the expansion of the ECEAP program to include full-day and extended day slots in 2014, and the passage of the Early Start Act in 2015 are explored in separate WSIPP reports and are not examined in this study.

[15] (DCYF). *2018-19 Outcomes Report* and (DCYF). *2019-20 Caseload Forecast Report*.

[16] Between 1985-2006, Part Day services included 240 hours of educational services over 30 weeks and each class session lasted at least 2.5 hours. Beginning in 2007, total educational hours increased from 240 to 320 hours over 30 weeks.

Exhibit 3

Key Changes, by Calendar Year



III. Data and Sample Construction

We rely on almost 20 years of administrative data from state agencies on ECEAP eligible children and their families. These include the following:

- Department of Children, Youth and Families (DCYF),
- Department of Social and Health Services (DSHS),
- Administrative Office of the Courts (AOC),
- Employment Security Department (ESD),
- Education Research Data Center (ERDC), and
- Office of Superintendent of Public Instruction (OSPI).

See Appendix I for a full description of data coverage by source.

Cohorts

We define cohorts using academic years (AY) running from September through August. This AY approach aligns with standard birthdate cutoffs for both ECEAP and kindergarten entry, the ECEAP academic year program, and subsequent outcome data.

Our analysis divides children into two distinct cohort groups.

1) Historical cohorts
2) Recent cohorts

Historical Cohort Group Construction

The historical cohort group includes all cohorts in the 2014 study[17] plus additional cohorts born in the three years prior. These children were born between September 1, 1996, and August 31, 2004. These children would have been eligible to participate in ECEAP as three-year-olds between 2001 and 2008 and as four-year-olds between 2002 and 2009.

The long follow-up period allows us to examine the impact of ECEAP on academic and behavioral outcomes through elementary school (short-term), and high school (long-term). All children in the historical cohort group participated in Part-Day classes, as ECEAP had not yet been expanded to offer School-Day or Working-Day models.

[17] Children included in the 2014 study were born between September 1, 1999, and August 31 2004.

Recent Cohort Group Construction

The recent cohort group includes all children who were born between Sept 1, 2004, and August 31, 2014. These children were born after those included in the 2014 study and would have been eligible to participate in ECEAP as three-year-olds between 2009 and 2018 and as four-year-olds between 2010 and 2019.

The children in this group participated recently enough that long-term outcomes are not yet available. In addition, children in this group participated during a period where many changes were made to the ECEAP program, including more mandated class hours for Part-Day classes, the offering of an expanded set of dosage options, and the implementation of Early Achievers. Most of the children in this cohort participated in Part-Day programming but some participated in School-Day or Working-Day classes. It is for these reasons that we analyze the impact of ECEAP for this group separately from the impact of ECEAP on the historical cohort.[18]

See Appendix II for a summary of the cohorts by age and grade level.

Identifying Treatment and Comparison Children

Our first step was to identify a sample of ECEAP eligible children. We focused on the income and age eligibility requirement since the vast majority of ECEAP participants meet these requirements.

Our primary goal in this evaluation is to identify and compare children who participated in ECEAP with similar children who would have been eligible for ECEAP but who did not participate. Since children whose families voluntarily applied for and attended ECEAP (the "treatment" group) might be systematically different from children whose families did not, we needed to find a dataset including information on key environmental, family, and other characteristics. Controlling for these characteristics increases our confidence that differences in outcomes are the result of ECEAP participation rather than other factors.

It was critical to match children on pre-treatment characteristics (before ECEAP participation) rather than on post-treatment characteristics (during/after ECEAP participation) because program participation may impact key variables of interest. For example, if participation in ECEAP allows parents to pursue educational or career opportunities, then matching on parental education or income indicators measured after ECEAP participation might obscure some of the program impacts because it could be partially determined by ECEAP participation.

[18] We do not explore the impact of program changes on children in recent cohort in this report. However, the impact of changes in ECEAP dosage on children's outcomes is explored in Cramer & Rashid (2022).
The impact of the introduction of the Early Achievers is explored in Goodvin, R., & Hansen, J. (2019). *Early Achievers evaluation report one: Background and research design* (Doc. No. 19-12-2202). Olympia, Washington State Institute for Public Policy; Goodvin, R., Rashid, A., & He, L. (2020). *Early Achievers evaluation report two: Prekindergarten quality and child outcomes in kindergarten* (Doc. No. 20-12-2203). Olympia: Washington State Institute for Public Policy; and Rashid, A., Goodvin, R., & Krnacik, K. (2021). *Early Achievers evaluation report three: Variation in links between quality and kindergarten readiness for children with childcare subsidy* (Doc. No. 21-12-2201). Olympia: Washington State Institute for Public Policy.

Identifying pre-ECEAP indicators was a major challenge because most data sources that might provide income and other information on children are collected after they would have participated in the ECEAP program (e.g., school records from OSPI).

To address this challenge, we received data from Research and Data Analysis (RDA)[19] on all children receiving DSHS services from either the Economic Services Administration (public assistance or childcare subsidy) or the Developmental Disabilities Administration (disability services) in the 20 months prior to the start of the year that they would have been eligible to enroll in ECEAP.[20]

Identifying Children with Requisite Household Level Information

We refined this dataset to include children who had requisite individual and household level information during the ECEAP eligibility determination period so that we can control for key characteristics. Included children had to receive Basic Food benefits, live with a parent, and meet other requirements discussed later in this section.

About a quarter of children in the DSHS services file were excluded at this step.

Received Basic Food Benefits. To be included in the study, ECEAP-eligible children had to be part of a household that received Basic Food benefits in the eligibility period.[21]

We restricted our sample to only include children receiving Basic Food benefits for two reasons. First, the income cutoff for Basic Food eligibility is similar to the income cutoff for ECEAP eligibility.[22] Second, administrative data on Basic Food receipt has the added benefit of including extensive demographic information on all other household members,[23] in addition to information on each ECEAP-eligible child. This assured us that we had information on the child, at least one parent, and their household members prior to the ECEAP program year.

In addition to this pre-ECEAP requirement, we further limited the sample to those who had received at least 12 months total of Basic Food services over the ECEAP program year and prior 20 months. See Exhibit 4 for an example.

[19] RDA pulled data for this assignment from the department of Social Health and Health Services' (DSHS) Integrated Client Databases. The data system contains more than two decades of information from several state agencies.

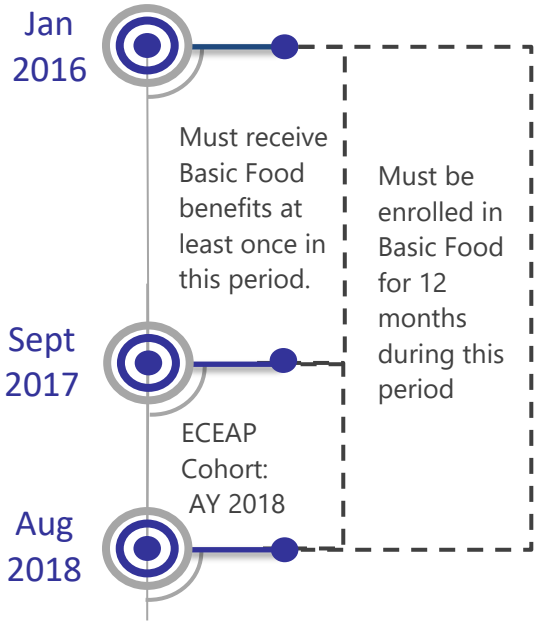
[20] For example, a child born on September 1, 1999, would have been eligible to attend ECEAP when they were three in the 2004 academic year and when they were four in the 2005 academic year. They would have been included in the 2004 sample if they received DSHS services from 1/1/2002 through 8/31/2003. They would have been included in the 2005 sample if they received DSHS services from 1/1/2003 through 8/31/2004. This eligibility definition was selected because it covered the many ways that children could prove their income eligibility for ECEAP enrollment.
Although this is similar to the data requested in the 2014 study, the data pull was new for the 2021 report. It included more information about participants and their families and included three additional birth cohorts of children. Differences between the 2014 data pull and the 2021 data pull are discussed in the report and the Appendix VI.

[21] The 2014 report also restricted their sample to only include children receiving Basic Food services for similar reasons.

[22] The income cutoff for ECEAP was 110% FPL for the historical and recent cohorts. The income cutoff for Basic Food was 130% for the historical cohort. However, on October 1, 2008, the Gross Income Standard for the Basic Food program was raised from 130% to 200% of FPL for Categorically Eligible (CE) households.

[23] Households are defined in this report as assistance unit members reported as part of a Basic Food services application. Please see WAC 388-408-0035 for a full description of who is included in Basic Food assistance units.

Exhibit 4
Example of Basic Food Enrollment Requirements for Study Inclusion
(For a child eligible to participate in the AY 2018 ECEAP cohort)



Children could appear in an ECEAP program year as a three- or four-year-old. We found that most children in our sample were eligible for ECEAP as a three-year-old and subsequently eligible as a four-year-old. To avoid double-counting children and to ensure that children were properly matched, we used information for the first eligibility period when children appeared in multiple years. Please see Appendix IV for more information.

Lived with At Least One Parent. Parental/ guardian characteristics can play an essential role in the decision to enroll a child in a program like ECEAP. We further restrict our sample to only include children for whom the household head is either a parent, co-parent, or stepparent to be able to control for parent characteristics in our analysis.

Children whose head of the household was a grandparent or other type of non-parent relationship (less than 5% of the total eligible child population) were excluded because it was difficult to determine guardianship responsibilities for other types of heads of households and control for parental characteristics.

Foster children were not specifically excluded, but they are likely to be underrepresented in the data because of the way that we constructed the sample.

Limitations with Identification Strategy. Our choice to limit our sample to children receiving Basic Food benefits and living with parents excludes some of the children that ECEAP was designed to support with their comprehensive services.[24] These excluded children may have different outcomes than children included in the sample.

We acknowledge that this reduces the generalizability of our analysis and excludes children who may have the most need for additional supportive services and the greatest potential benefit from participation in an early childhood education program. However, we are not able to reliably capture the information needed to include these groups in our analysis.

[24] Research by DCYF found that children in foster care were less likely to meet WA Kids assessment standards or 3rd grade assessment standards than children in the general population. (See: Education Outcomes of Children and Youth Experiencing Foster Care).

Identifying ECEAP Participants

Once we identified ECEAP-eligible children in the Basic Food data sample, we identified which children attended ECEAP (treatment) and which did not (comparison). We use enrollment data to identify all children who were ever enrolled in ECEAP[25] in either the historical or recent cohorts.

We found that 90% of children who had the requisite individual information could be identified in some of the K-12 data and ECEAP enrollment status could be determined.

Of the ECEAP participants, 40% had requisite individual and household information in the Basic Food data. While this is comparable to the percentage in our 2014 study, a majority of ECEAP participants, and comparable children, are not captured. While this gap was larger than expected, we did not expect to have full coverage of ECEAP participation in the Basic Food sample.[26]

Additionally, exclusions of data from the analysis due to quality and completeness issues are expected when using administrative data from real-world program implementation for program evaluation.

While some generalizability may be lost due to these exclusions, and we cannot comprehensively capture outcomes for all ECEAP participants, we are able to create the most reliable picture of outcomes for ECEAP's participants versus outcomes for similar children who were eligible for ECEAP but did not participate by restricting the dataset to include children with the most complete pre-treatment characteristic data available.

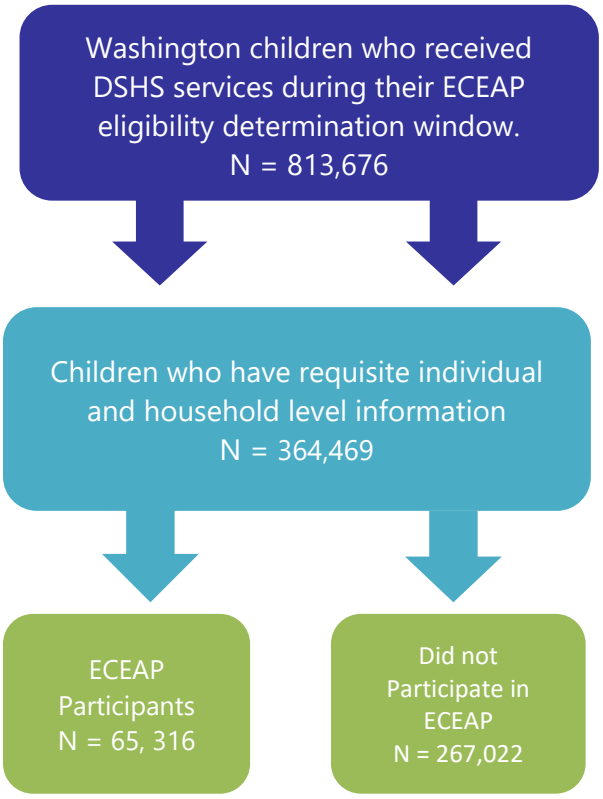
Exhibit 5 outlines the RDA sample size at each of these stages.[27] For more information about the difference between the samples, please see Appendix III.

[25] Our main analysis does not account for intensity of the ECEAP treatment—that is, a child attending ECEAP for a partial year is treated the same as a child attending ECEAP for the entire year. Similarly, this variable does not account for children who have attended ECEAP for more than a year. Considering all children to be ECEAP participants regardless of intensity of their treatment is more consistent with estimating an "intent to treat" effect. As a sensitivity test, we restrict the sample to only include children who were enrolled in ECEAP for at least 180 days in the Appendix.

[26] Patton, D., Liu, Q., & Felver, B.E.M. (2018). *Service Use, Risk Factors, and Assessments among ECLIPSE, ECEAP and ESIT Clients* found that 88% of ECEAP participants between SFY 2013 and SFY 2016 received Basic Food during that same period. This number represents and upper bound of overlap. Our approach adds restrictions for inclusion, such as requiring a minimum of 12 months of Basic Food enrollment. It also restricts the time frame during which children must receive services. We require children to receive services at least once in the 20 months prior to potential ECEAP enrollment, while the 2018 report used a four-year inclusion period. Patton, D., Liu, Q., & Felver, B.E.M. (2018). ***Service use, risk factors, and assessments among ECLIPSE, ECEAP, and ESIT Clients***. Department of Social & Health Services. Olympia: WA.

[27] It does not display the reduction in the ECEAP sample. For more information, please see the ECEAP data processing section in Appendix III.

Exhibit 5
Identification of Program and Comparison Groups



Note:
Includes historical and recent cohorts.

Identifying Children with Outcome Information

Finally, we determined if the children appeared in the outcome data of interest. We are unable to include children who either did not enroll in Washington public schools or do not have a record of the outcome of interest in the data provided. The exact match depends on the outcome examined.

IV. Evaluation Methodology

In an ideal world, the impact of ECEAP would be explored by randomly assigning eligible children to ECEAP or a control condition and then observing later education outcomes. Since this was not possible for this assignment, we match children who participated in the program to comparable children.

Key Differences Between ECEAP Participants and Non-Participants

There are many differences between children who did and did not participate in ECEAP prior to matching. Due to the large sample size, some of these differences are statistically significant, but perhaps not practically meaningful.

ECEAP participants are more likely to be Black or Hispanic. They are also more likely to speak Spanish. They are less likely to receive Working Connections Child Care subsidies (WCCC). They are more likely to have an out-of-home placement reported in the previous year.

The head of household is more likely to have less than a high school diploma and more likely to be unemployed.

The household income (measured as the calculated percent of poverty level) is lower for ECEAP participants than non-ECEAP participants.

Methodology

To address the underlying differences between the treatment and comparison groups, we use Coarsened Exact Matching (CEM) to match ECEAP children to comparable non-participants. Children in the historical cohort group are matched and analyzed separately from children in the recent cohort group.

To perform CEM, we identify a subset of characteristics we believe may influence both the decision to enroll in ECEAP and later outcomes. We then try to match ECEAP participants to children in the comparison group.

Binary characteristics (e.g., sex), do not need any further manipulation for CEM. For continuous and some categorical characteristics (e.g., years of education), it may be necessary to "coarsen" matches, or consolidate granular data into broader categories on which to match children. For example, in our analysis, we were unable to match children on exact years of parental education. We converted the original years of education data into a new variable with three bins—1) parental education is less than high school graduate, 2) high school graduate or equivalent, 3) some college or more.

ECEAP participants are then matched to children who share the same characteristics across all variables used in the matching (they will have the same age, race, parental years of education, etc.). Children who do not have matches are dropped from the sample.

After matching, we run a regression modeling the effect of ECEAP controlling for the variables included in the CEM, additional child, parent, and household characteristics on which we were unable to match, and school characteristics (when applicable) while adjusting for weights generated by the coarsened exact matching. We used linear regression for continuous outcomes (e.g., math assessment) and a logistic regression for binary outcomes (e.g., high school graduation).

See Appendix IV for a more in-depth discussion of the methodology and analysis decisions.

Change From 2014 Methodology

The use of CEM is a departure from the methodology used in our 2014 report.

The 2014 WSIPP report used distance (as the crow flies) between the child's census tract of residency and closest ECEAP center as an instrumental variable (a variable that acts as a proxy for quasi-random treatment assignment) for ECEAP participation. The intuition behind this decision was that the distance from an ECEAP center might be related to the decision to enroll in ECEAP but not to other selection factors (e.g., income, or parent's education), and therefore would make a good proxy to random treatment assignment.

The change in methodology resulted in different findings. Specifically, while the 2014 report found a positive, statistically significant association between ECEAP and test scores, we do not find a positive, statistically significant relationship between test scores and ECEAP participation. A full discussion of the changes to our methodology and findings is provided in Appendix VI

Limitations of Methodology

While we believe that the results from the CEM most transparently represent the relationship between ECEAP participation and later outcomes, we must note that this method is susceptible to certain limitations. While CEM allows us to observe the differences in outcomes between ECEAP participants and similar non-participants, it does not allow us to know if ECEAP participation itself caused those differences.

Despite the extensive amount of administrative data provided for sample children, we are only able to control for a small fraction of what goes into the choice of enrolling a child in ECEAP, or any ECE program. As a result, some of the observed differences could be due to omitted factors impacting both the decision to enroll in ECEAP and the measured outcome. We would then falsely be attributing a portion of the impact of these omitted factors to ECEAP participation.

For example, a parent that highly values academics and academic preparedness may choose to enroll their child in ECEAP. This parent may also make other choices to ensure their child's academic success. This child may have better standardized test scores, even if they had not participated in the ECEAP program, because of their parent's prioritization of academic success. If we are unable to control for the parent's values, we will overcount the contribution of ECEAP to the child's later success. Conversely, ECEAP sites may specifically select children who are the most at risk, controlling for observed demographic information. These children may have had worse outcomes than the comparison group had they not participated in the ECEAP program. We would undercount the contribution of ECEAP to the child's later success by not being able to control for the unobserved vulnerability.

Related to this concern, it is important to note that when we improve the quality of the match of the comparison group to ECEAP children through CEM, we may increase the likelihood that children in the comparison group are participating in some other, unobserved ECE program, like Head Start. Having alternative treatments in the comparison group creates the possibility for multiple underlying reasons for the results.

A finding of a statistically insignificant or negative impact does not necessarily imply that ECAEP is ineffective. Since "treatment as usual" may include other, unobservable ECE programs (such as Head Start), the comparison is not ECE participants vs non-ECE participants. Instead, this is a comparison of ECEAP participants with other, similar children who may be eligible for, and enrolled in, other ECE programs. With the data available, we are unable to answer the more specific question of whether ECEAP improves outcomes for children compared to children without access to ECE.

For the reasons outlined above, we caution against interpreting the results as the causal impact of ECEAP. However, our findings capture the relationship between ECEAP with later outcomes as best as possible with the available data.

V. Results

This evaluation adds to the existing research by reporting on the following outcomes in the context of Washington State:

- Academic outcomes,
- Behavioral outcomes, and
- Contemporaneous family outcomes.

Exhibit 6 (on the next page) summarizes all outcomes included in our study as well as the cohorts for which the analysis is available.[28]

[28] Complete results for all outcomes can be found in Exhibits A13-A20 in Appendix V.

Exhibit 6

Outcome Measures and Outcome Availability by Cohort

Outcome	Description	Population where the outcome is available	
		Historical group	Recent group
Academic outcomes			
Kindergarten readiness: Overall	Whether the child was deemed kindergarten ready on 6 of 6 domains of the Washington Kindergarten Inventory of Developing Skills (WaKIDS),		X
Kindergarten readiness: By domain	Whether the child was deemed kindergarten ready on each domain of the Washington Kindergarten Inventory of Developing Skills (WaKIDS),		X
Assessment: Math	Standardized test score (z-score) on the WASL, MSP/HSPE, or SBA math assessment by grade.	X	X
Assessment: Reading	Standardized test score (z-score) on the WASL, MSP/HSPE, reading assessment by grade.	X	
Assessment: Writing	Standardized test score (z-score) on the WASL, MSP/HSPE, writing assessment by grade.	X	
Assessment: ELA	Standardized test score (z-score) on the SBA English language arts/literacy assessment by grade.		X
Special education K-8	Indicator of participation in special education, by grade	X	X
Special education 9-12	Indicator of participation in special education, by grade	X	
High school graduation	Indicator of on-time graduation as measured by graduating by the expected graduation year reported in their first appearance in 9th grade (or first high school year)	X	
Behavioral outcomes			
Any misdemeanor	Indicator of any misdemeanor conviction recorded by Administrative Office of the Courts (AOC), by age	X	
Any felony	Indicator of any felony conviction recorded by AOC, by age	X	
Any juvenile justice conviction	Indicator of any conviction recorded by AOC, by age	X	
Family outcomes			
Parental hours worked	The difference in hours worked between the ECEAP academic year and the prior academic year	X	X
Parental wages earned	The difference in wages earned between the ECEAP academic year and the prior academic year	X	X
CPS involvement	Indicator of whether the family had any CPS-investigation or CPS- Family Assessment Response (FAR) during a potential ECEAP enrollment period (did not have to involve the ECEAP eligible child).	X	X

Notes:
The binary outcomes are *italicized*.
Family outcomes are based on individuals who were eligible to participate in ECEAP at age four. This excludes individuals who participated when they were three and individuals who were only eligible to participate when they were three.
For more information, please see Appendix IV.

For all outcomes, we report outcomes separately for individuals in the historical cohort group and individuals in the recent cohort group.

For all binary outcomes, we report the outcomes for the ECEAP group and non-ECEAP group, the percentage difference between the two, and the level of statistical significance.

Exhibits 7-13 highlight our key results. Additional analysis is included in the appendix.

All outcomes also report the p-value associated with the outcome. The p-value can range from 0 to 1 and represents the chance that we would observe the reported effect if the intervention truly had no effect at all. P-values closer to zero represent observations less likely to occur by chance.

- Differences that are significant at the 0.01-level are indicated with ***
- Differences that are statistically significant at the 0.05-level are indicated with **
- Differences that are statistically significant at the 0.10-level are indicated with *

Academic Outcomes

For our academic outcomes, we examine the relationship between ECEAP and kindergarten readiness, 3rd- through 5th-grade test scores, special education participation, and high school graduation.

Kindergarten Readiness

Kindergarten readiness is defined as meeting/exceeding expectations in all six kindergarten readiness domains on the Washington Kindergarten Inventory of Developing Skills (WaKIDS) assessments. [29] We also report the relationship between ECEAP and meeting/exceeding expectations on each of the six domains.

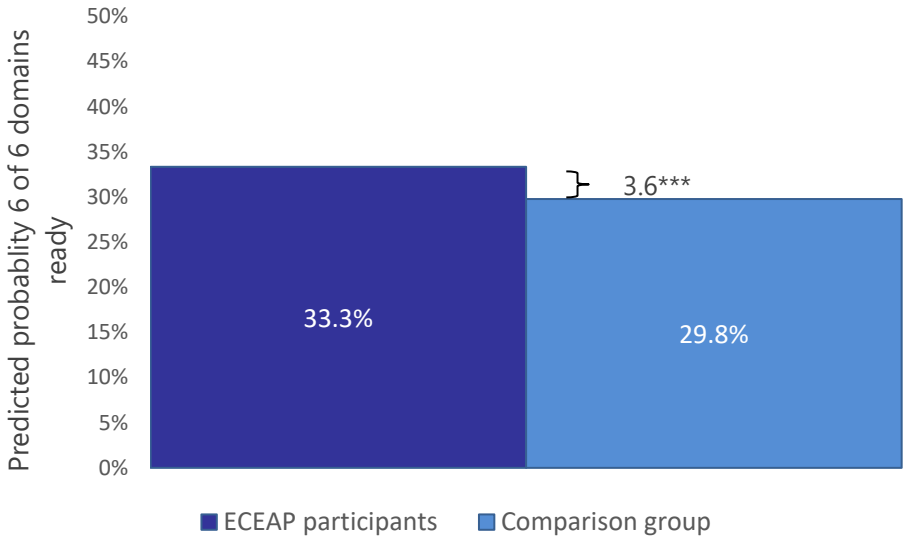
Our results only capture kindergarten readiness for the subset of children who took the assessment from the 2016 through the 2020 academic years.[30] Children who received these assessments would have participated in ECEAP from the 2014[31] through the 2019 academic year so they only represent part of the recent cohort group.

We find that participation in ECEAP is positively associated with performance on WaKIDS. This result was statistically significant. Children who participated in ECEAP were approximately 12% more likely to be kindergarten-ready in all six WaKIDS domains than non-participants (Exhibit 7). The difference between ECEAP participants and non-participants was positive on all domains (Exhibit 8). The difference was largest for literacy and smallest for social-emotional skills.

[29] ERDC website: *Early learning feedback report.*
[30] While the WaKIDS assessments started in 2012, the assessment objectives and dimensions have only been consistent from AY 2016 through AY 2020. The last cohort included children who took the assessment in fall 2019, therefore these results only speak to outcomes pre-COVID.
[31] This would not include children who participated in ECEAP as four-year-olds in the 2014 academic year.

Exhibit 7

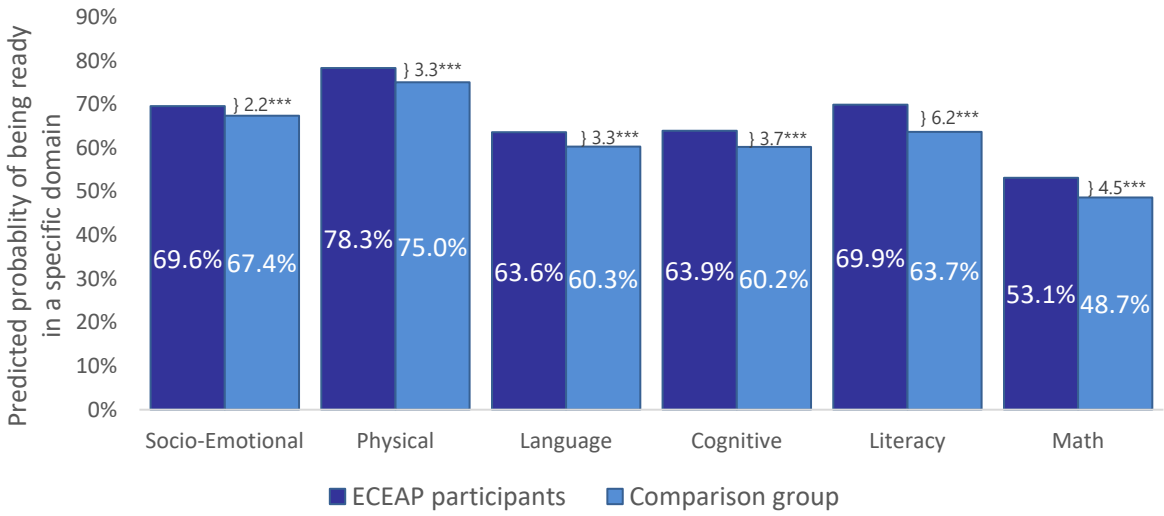
Predicted Rates of Being Kindergarten-Ready in 6 of 6 WaKIDS Domains—Recent Cohort



Notes:
Predicted probabilities are estimated from a single logistic regression model using the full set of control variables listed in Appendix IV while adjusting for weights generated by the coarsened exact matching.
*** significant at the 0.001-level, ** significant at the 0.05-level, * significant at the 0.10-level.
The percentage point difference in the probability of kindergarten readiness across the two groups is 3.6% = 33.3% - 29.8%. The percent difference is 12.0% = (33.3% - 29.8%) ÷ 29.8%.

Exhibit 8

Predicted Rates of Being Kindergarten-Ready in each WaKIDS Domain—Recent Cohort



Notes:
Predicted probabilities are estimated from a single logistic regression model for each domain using the full set of control variables listed in Appendix IV while adjusting for weights generated by the coarsened exact matching.
*** significant at the 0.001-level, ** significant at the 0.05-level, * significant at the 0.10-level.
The percentage point difference in the probability of kindergarten readiness across the two groups ranges from 2.2 to 6.2 percentage points. The percent difference ranges from 3.2% to 9.8%.

Test Scores

We use data on the Washington State-administered standardized test scores in 3rd through 5th grade by subject area.[32] These are the same test grades used in our 2014 report for the historical cohort group. These results are meant to duplicate the analysis done in 2014 to determine if there are any differences in test score results from using a new statistical methodology. We also report 3rd-, 4th-, and 5th-grade test score outcomes for the recent cohort group.

Assessments change over time and across grades; they cannot be directly compared in their raw form. To make them comparable within a cohort and across grades we converted test scores into z-scores. The z-score standardizes the test score for each test type, year, subject, and grade and will have a mean score of zero and a standard deviation of one for the entire state of Washington. For the historical and recent cohort groups, the mean z-scores were negative, which means that the average test scores for the students included in our sample were lower than the statewide average.

Historical Cohort Group Test Scores. We find that ECEAP participants' 3rd-grade math z-scores were higher than children in the comparison group (-0.432 and -0.458, respectively). This means that ECEAP participants' scores were 0.026 standard deviations higher than the comparison group. These results were small but statistically significant. We do not find evidence of a positive statistically significant relationship between ECEAP and test scores for any other grade level/subject pair in the historical cohort group.

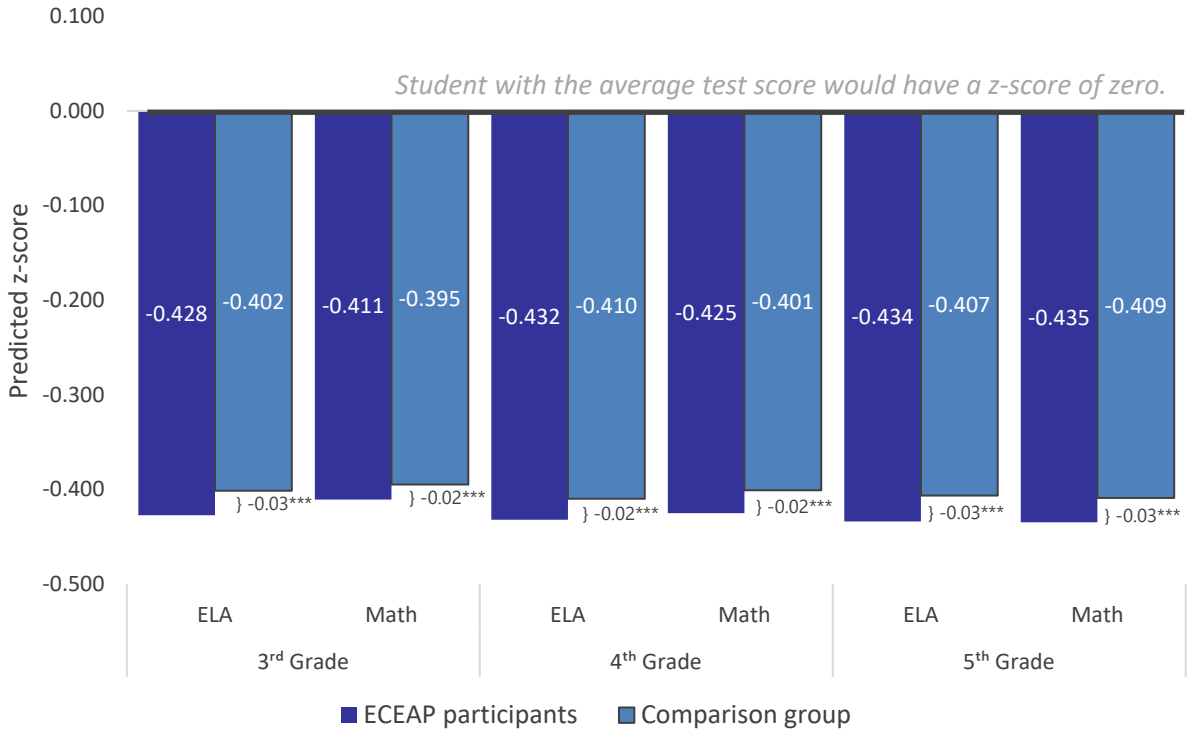
This runs counter to our 2014 results, where we found the difference in test scores was positive and statistically significant for all grades and subject areas. It seems that the difference between the 2014 and current impact on test scores is largely due to the changes in methodology. Please see Appendix VI for a full discussion.

Recent Cohort Group Test Scores. For the recent cohort group, we find that the difference between test scores for ECEAP and the comparison group is negative and statistically significant for all subjects and grade levels analyzed. These negative impacts are small relative to the difference in performance of the population compared to the general student population (Exhibit 9).

[32] Three different standardized tests were used during the analysis years. Children in the historical cohorts took either the Washington Assessment of Student Learning (WASL) (administered between 2006-2009) or the Measurements of Student Progress (MSP) (administered between 2010-2014). Assessments for the recent cohort groups are based on the Smarter Balanced Assessment (SBA) (administered between 2015 onwards). These results exclude children who took the basic version of these test. We a separate analysis with these scores included and it did not meaningfully change the results. When Washington switched from the MSP to the SBA it also switched from testing children on reading in 3rd, 4th and 5th grade and writing in 4th grade, to testing ELA in 3rd, 4th, and 5th grade. Test scores were converted to z-scores so that we could combine years where different tests were used in the analysis

Exhibit 9

Predicted Test Score Relative to Standardized State Mean—Recent Cohort



Notes:
Lower, negative, predicted z-scores indicate a lower test score and worse performance on the assessment. For example, ECEAP participants performed 0.428 standard deviations lower than the average student performance on 3rd grade ELA scores, this is slightly worse than the comparison group, who performed 0.402 standard deviations below the average.
Predicted z-scores are estimated from a single linear regression model for each grade and subject using the full set of control variables listed in Appendix IV while adjusting for weights generated by the coarsened exact matching.
*** significant at the 0.001-level, ** significant at the 0.05-level, * significant at the 0.10-level.

Special Education

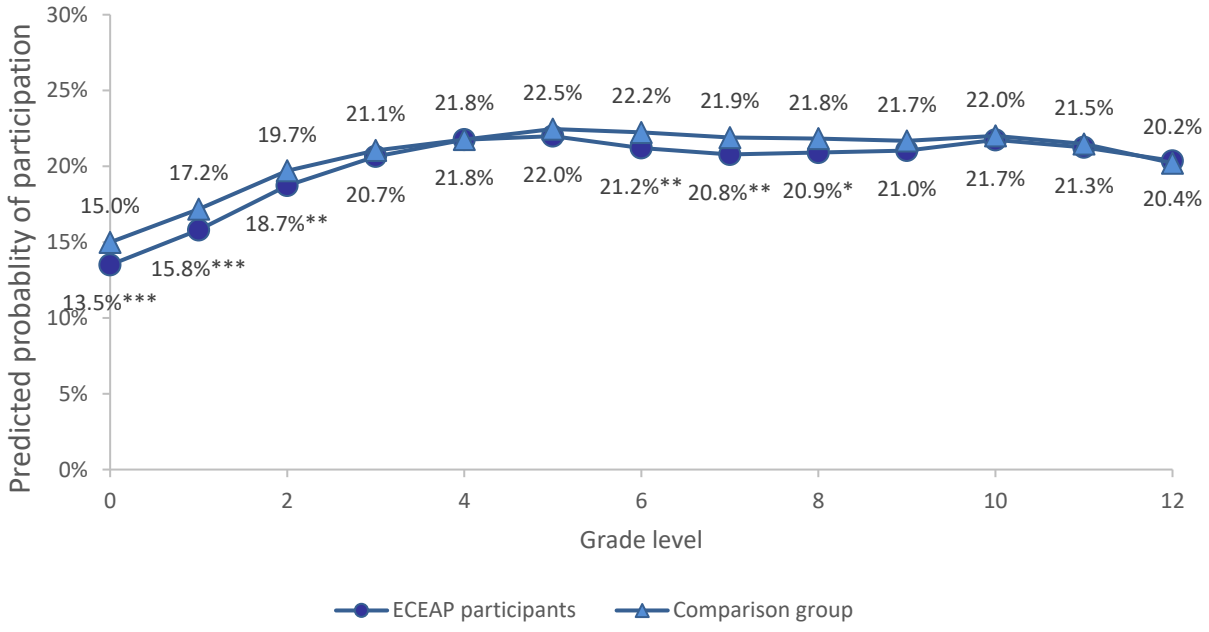
We examine the association between ECEAP participation and enrollment in special education by grade, for both the historical and recent cohort groups.

We find that in the historical cohort group, children who participate in ECEAP are slightly less likely to participate in special education than children in the comparison group over their K-12 schooling years. These differences are only statistically significant in grades K-2, 6th grade, and 7th grade, as shown in Exhibit 10. However, taking kindergarten as an example, a 1.4 *percentage point* difference in special education enrollment represents a 10% reduced likelihood of special education participation for ECEAP participants relative to the comparison group.

For the recent cohorts, we observe a smaller, but more persistent difference between ECEAP participation and special education participation by grade (Exhibit 11). For these cohorts, the special education participation rate is usually under 1 percentage point lower for ECEAP participants. However, the difference is statistically significant for 1st through 5th grade.

Exhibit 10

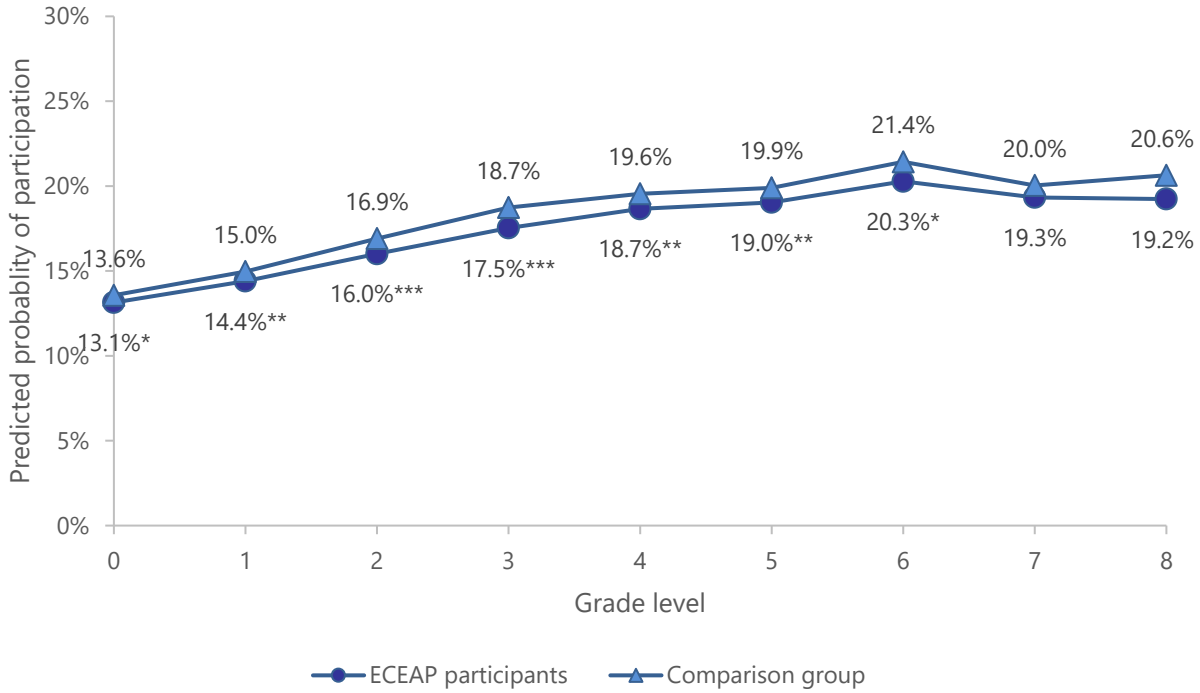
Predicted Rates of Participating in Special Education in Each Grade Level—Historical Cohort



Notes:
Predicted probabilities are estimated from a single logistic regression model for each grade using the full set of control variables listed in Appendix IV while adjusting for weights generated by the coarsened exact matching.
*** significant at the 0.001-level, ** significant at the 0.05-level, * significant at the 0.10-level.

Exhibit 11

Predicted Rates of Participating in Special Education in Each Grade Level—Recent Cohort



Notes:
Predicted probabilities are estimated from a single logistic regression model for each grade using the full set of control variables listed in Appendix IV while adjusting for weights generated by the coarsened exact matching.
*** significant at the 0.001-level, ** significant at the 0.05-level, * significant at the 0.10-level.

High School Graduation

We estimate the association between ECEAP participation and on-time high school graduation for most children in the historical cohort group.[33] We are able to observe this outcome for two of the three cohorts included in the original analysis in addition to three cohorts who participated in ECEAP prior to those included in our original analysis.

We find the likelihood of graduating from high school is 67.6% for ECEAP participants compared to 66.9% for non-participants. This 1% difference is not statistically significant.

Behavioral Outcomes

For behavioral outcomes, we examined the association between ECEAP participation and juvenile convictions for the historical cohort group.[34]

We begin counting convictions in the academic year a child turns 12 through the end of the academic year a child turned 14, 15, 16, or 17.[35] Ultimately, we treat this outcome as a binary variable rather than a count of convictions; this means the outcome can be interpreted as "at least one conviction by the end of the academic year they turn X years of age." As age increases, our sample size expectedly decreases because we have less follow-up data for younger cohorts included in our sample. We find that the difference between the likelihood for conviction between ECEAP participants and non-participants was less than 1% and not statistically significant for any conviction type (misdemeanor, felony, or any) at any age. For example, by the end of the academic year a child turned 17, roughly 13% of youth in both groups had ever been convicted of any crime; just under 5% of youth in both groups had been convicted of a felony. For more details, please see the Appendix.

[33] We define on time high school graduation as graduating by the expected graduation date first reported when the student first appears in 9th-grade or higher. This is equivalent to a graduating within four years of 9th-grade enrollment.
[34] The recent cohort group is too young to have meaningful juvenile convictions information.

[35] We used 12 as a cutoff because children are unlikely to be tried for committing a crime before they are 12. See RCW 9A.04.050. Results for age 14, 15, and 16 are included in the Appendix.

Contemporaneous Family Outcomes

For family outcomes, we examined the association between ECEAP participation and events that occurred during the academic year of ECEAP, for both the historical and recent cohort groups. Specifically, our outcomes were differences in hours worked and in wages earned by the head of household during the academic ECEAP year compared to the prior academic year. We were interested in whether participation in ECEAP allowed for more hours of work or potentially higher wages.

We also looked at household Child Protective Service (CPS) involvement during the ECEAP academic year. A household is defined as having a CPS involvement if the assistance unit had any CPS-investigation or CPS-Family Assessment Response (FAR) [36] during a potential ECEAP enrollment period (the event did not have to involve the ECEAP-eligible child). [37]

Parental Employment

We found that heads of households increased their hours worked and earned higher wages in both the ECEAP and comparison groups. There were no significant differences between groups with regards to the increased hours worked.

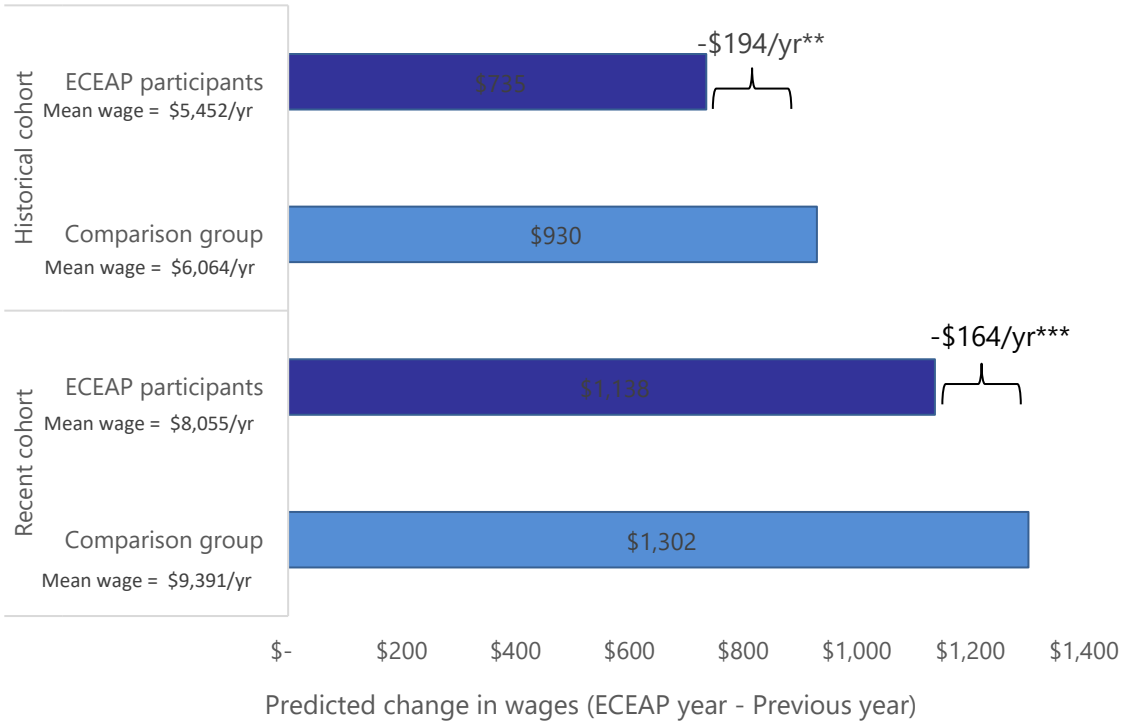
We found gains in wages earned were lower for ECEAP parents compared to non-ECEAP parents (Exhibit 12). In the historical cohort parents of ECEAP children saw an increase in wages by $735/year and the comparison group saw an average increase of $930/year. In the recent cohort parents of ECEAP children saw an increase in wages by $1,138/year and the comparison group saw an average increase of $1,302/year. Parental earnings are not adjusted for inflation because it will not impact the difference in earnings between the treatment and comparison groups, after matching. However, this may partially explain the difference in earning impacts between the two cohort groups. These differences in earnings gained between ECEAP parents and the comparison group are significant but amount to less than $200 per year for both cohorts.

[36] FAR is a CPS alternative response to a screened-in allegation of abuse or neglect. FAR focuses on children and youth safety along with the integrity and preservation of families when lower risk allegations of maltreatment have been screened-in for intervention.

[37] We also examined the relationship between ECEAP enrollment and CPS enrollment involving the ECEAP eligible child. We found that expanding cases to the entire household did not change the results.

Exhibit 12

Predicted Change in Wages Earned per Year during ECEAP Program Year, by Cohort



Notes:
Predicted change in wages earned is estimated from a single linear regression model for each cohort using the full set of control variables listed in Appendix IV while adjusting for weights generated by the coarsened exact matching.
*** significant at the 0.001-level, ** significant at the 0.05-level, * significant at the 0.10-level. The difference in the predicted *extra* wages earned across the two groups is -$194 per year for the historical cohort, and -$164 per year for the recent cohort. The percent difference in *extra* wages earned per year between the two groups is -20.9% for the historical cohort and -12.6% for the recent cohort.

CPS Involvement

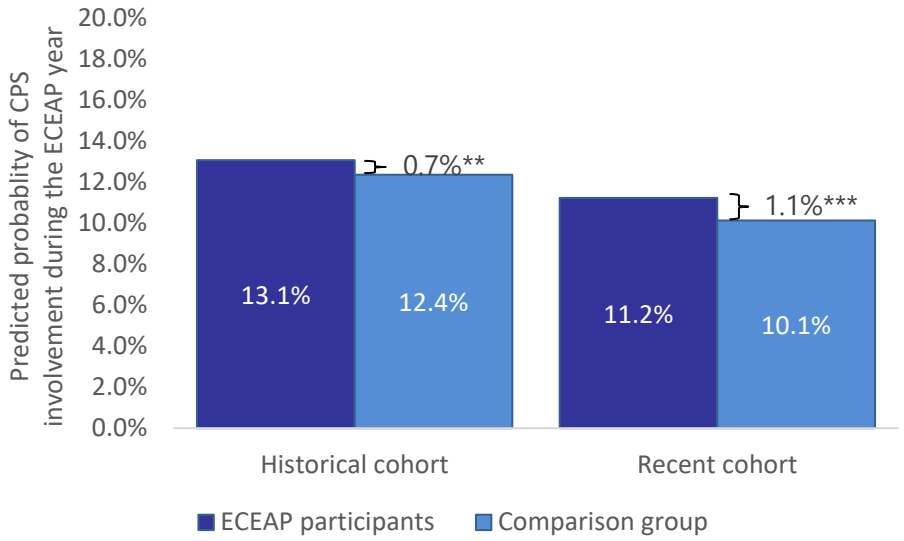
There was a small, albeit statistically significant difference in CPS involvement between ECEAP participants and the comparison group in the recent cohort. ECEAP participants had a higher rate of CPS involvement by 0.7 percentage points in the historical cohort and 1.1 percentage points in the recent cohort. This translates to a 5.7% and 10.9% increase in CPS involvement for each cohort group respectively (Exhibit 13).

The meaning of increased CPS involvement is a bit ambiguous. Although it is possible that ECEAP is causing more household instability or that households in the comparison condition are receiving a more effective, unobserved, intervention, it is also possible that suspected abuse and neglect are reported more frequently for the children in ECEAP.

ECEAP staff, as well as staff at daycare centers and preschools, are "mandatory reporters." That is, they are legally obligated to report suspected abuse or neglect. Since some of the children in the comparison group were probably not in preschool, they may have had less contact with mandatory reporters. Given the nature of the comprehensive services provided by the program, ECEAP sites may be better able to identify when a household may need CPS involvement.

Exhibit 13

Predicted Rates of CPS Involvement During ECEAP Program Year, by Cohort



Notes:
Predicted probabilities are estimated from a single logistic regression model for each cohort using the full set of control variables listed in Appendix IV while adjusting for weights generated by the coarsened exact matching.
*** significant at the 0.001-level, ** significant at the 0.05-level, * significant at the 0.10-level.
The difference in the probability of CPS involvement across the two groups is 0.7 percentage points for the historical cohort and 1.1 percentage points for the recent cohort. The percent difference between the two groups is 5.7% for the historical cohort and 10.9% for the recent cohort.

VI. Summary and Limitations

Overall, we find that ECEAP participants have modestly better academic outcomes. We do not find a significant difference in behavioral outcomes for participants compared to non-participants.

Although several of the family outcomes were statistically significant, the practical significance and/or meaning of the findings were ambiguous.

Exhibit 14

Summary of Report Outcomes

Outcomes	Historical	Recent
Academic outcomes		
Kindergarten readiness: Overall	-	**Increases**
Kindergarten readiness: by Domain	-	**Increases**
Assessment: Math	**Mixed** (Statistically significant increase in the 3rd grade only)	**Decreases**
Assessment: Reading	Not statistically significant	-
Assessment: Writing	Not statistically significant	-
Assessment: ELA	-	**Decreases**
Special education	**Decreases** (Statistical significance varies by grade)	**Decreases** (Statistical significance varies by grade)
High school graduation	Not statistically significant	-
Behavioral outcomes		
Any misdemeanor	Not statistically significant	-
Any felony	Not statistically significant	-
Any juvenile justice conviction	Not statistically significant	-
Family outcomes		
Parental hours worked	Not statistically significant	Not statistically significant
Parental wages earned	**Decreases**	**Decreases**
CPS involvement	**Increases**	**Increases**

The difference between ECEAP participants and non-participants appears to be largest immediately after ECEAP participation and (generally) fades over time. For example, ECEAP has a clear impact on kindergarten readiness and Special Education participation in early school years, while later school years and long-term outcomes have a more mixed finding. This result is consistent with recent literature on the impact of ECE programs, including our meta-analysis.

We also found that our results are not very sensitive to the choice of method used in our analysis (please see Appendix V for further discussion).

Limitations

As we stated previously, we are not able to make a causal assertion in our analysis, meaning we cannot conclude that the differences we observed between ECEAP participants, and the comparison group are *because* of ECEAP participation. Many factors impact the likelihood of ECEAP participation and later performance that cannot be controlled for in a retrospective analysis.

Families that choose to enroll their children in ECEAP may be different from families who do not participate. If parents who are the most motivated to ensure that their child succeed are taking advantage of the ECEAP program, then these kids might be Kindergarten ready even in the absence of this program. Conversely, if ECEAP enrolls children who will fare worse than their peers in the absence of an intervention, after controlling for observable characteristics, then the impact of ECEAP in this report will be smaller than the true benefit of the program.

In addition, we do not have information on the childcare or ECE participation for children in the comparison group. For example, we know that some children participate in private preschool programs or Head Start, but we are unable to observe that participation in the administrative data. Our results are showing the impact of ECEAP relative to all other options available, rather than no treatment. Results indicating that outcomes are not different for ECEAP participants compared to non-participants or even that ECEAP participants have less optimal expected outcomes than non-participants is not necessarily an indication that ECEAP is not working.

Finally, we used children already receiving DSHS services to create our sample. However, this does not fully capture the universe of children who may be eligible for ECEAP. Families must actively engage with Basic Food and other DSHS services. It is probable that families who receive services may be fundamentally different from families who do not receive services. These results may not generalize to eligible children not receiving DSHS services.

Future Work

This report examines the relations between ECEAP participation and subsequent outcomes for a subset of eligible children. We find the strongest relationship between ECEAP participation and early outcomes. Although we can describe the observed relationship between participation and outcomes, we cannot explain why we are seeing these results.

ECEAP-eligible children experience many interventions after preschool. These interventions are designed to help children build upon previous academic success or make up for deficits. Future research is needed to determine if the initial gains from ECEAP participation are dissipating or if the success of other K-12 programs is effectively helping children who were not exposed to ECEAP catch up. If the program effects are dissipating, it is important to understand why these children are not able to build on the initial gains. If other programs are effective at helping children who did not have the opportunity to participate in ECEAP catch-up, more research is needed to determine which programs are most effective at filling the gap. By evaluating the constellation of education programs, policymakers may be able to determine the optimal mix of interventions that will have the greatest overall impact on educational outcomes.

Acknowledgments

This report required a tremendous amount of data and support from several state agencies. We would like to thank all the staff at DCYF, DSHS, OFM, and OSPI who made this report possible. A special thanks to Thomas Aldrich, Bonnie Nelson, Tim Norris, Deleena Patton, Barbara Lucenko, Marci Arthur, Shane Riddle, Ezra Paskus, Vickie Ybarra, Sara Schwartz Jewell, Kevin Cummings, Karin Ganz, Sarah Veele, Angela Abrams, Markisha Lynch, Brandy Franco, Warren Wessling, Joyce Kilmer, and Lucas Snider. Thank you for your support in providing administrative data, thoughtful feedback, and patience with our many questions and revisions.

We would also like to thank our WSIPP colleagues including, Julia Cramer, Amani Rashid, Rebecca Goodvin and Marna Miller for being thought partners as we prepared this report.

Appendices

Evaluation of the Early Childhood Education and Assistance Program:
Short- and Long-Term Outcomes for Children

I. Summary of Data Received

Exhibit A1

Data Sources used for ECEAP Evaluation

Data type	Data systems or reports	Data sources/ agency providing data
Child demographics (e.g., birth month, birth year, races, sex)	ICDB Demographics	RDA
Child care subsidy participation	SSPS	DCYF/RDA
Child ECEAP eligibility and enrollments	EMS ELMS	DCYF/ERDC
Child academic program participation and assessment data	CSRS CEDARS	OSPI/ERDC
Child graduation data	CEDARS	OSPI/ERDC
Child juvenile justice contact		AOC/RDA
Parent demographics (e.g., age, race,)	ICDB Demographics	RDA
Parent level of education, marital status	ACES	ESA/RDA
Quarterly wages earned and hours worked	Unemployment Insurance Data	ESD/RDA
Number of individuals in the household	ACES	ESA/RDA
Household FPL* (Constructed from individual-level earned and unearned income information)	ACES	ESA/RDA
Relationship of AU members to the Head of Household	ACES	ESA/RDA
Any out of home placement or CPS involvement	Famlink	DCYF/RDA
Census tract (neighborhood-level) characteristics	2000 Census (Used for historical cohort geographic characteristics)	US Census Bureau via IPUMS
Census tract (neighborhood-level) characteristics	American Community Survey (Used for recent cohort geographic characteristics)	US Census Bureau via IPUMS
ECEAP site characteristics	EMS ELMS	DCYF/ERDC
School demographic information	Washington School Report Card	OSPI

Notes:

ACES: Automated Client Eligibility System
AOC: Administrative Office of the Courts
DCYF: Department of Children, Youth, and Families.
ELMS: Early Learning Management System.
EMS: Historical early learning management system
ERDC: Education Research & Data Center.
ESA: Economic Services Administration
ESD: Employment Security Department
SSPS: Social Service Payment System
CCA: Child Care Aware
CEDARS: Comprehensive Education Data and Research System
CSRS: Core Student Record System
OSPI: Office of the Superintendent of Public Instruction
RDA: Research and Data Analysis at the Department of Social and Health Services

II. Cohorts by Age and Grade Level

Exhibit A2

Cohorts by Age and Grade Level

		Age (by end of the academic year)																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		Grade (by end of the academic year)																	
		Born by Aug 31...				ECEAP (3)	ECEAP (4)	K	1	2	3	4	5	6	7	8	9	10	11	12
2014 study cohort	Historical cohort	**97**	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
		98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
		99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
		02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	
		03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19		
		04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19			
	Recent cohort	**05**	06	07	08	09	10	11	12	13	14	15	16	17	18	19				
		06	07	08	09	10	11	12	13	14	15	16	17	18	19					
		07	08	09	10	11	12	13	14	15	16	17	18	19						
		08	09	10	11	12	13	14	15	16	17	18	19							
		09	10	11	12	13	14	15	16	17	18	19								
		10	11	12	13	14	15	16	17	18	19									
		11	12	13	14	15	16	17	18	19										
		12	13	14	15	16	17	18	19											
		13	14	15	16	17	18	19												
		14	15	16	17	18	19													

Note:
The table shows the last two digits of the academic year.

III. Sample Construction

Data Acquisition[38]

We received comprehensive data from DSHS's RDA on all children receiving DSHS's services that were age-eligible for ECEAP for our years of interest (2002-2019) (N=813,676).

We reduced this sample to only include ECEAP eligible children who were members of households receiving Basic Food (N=664,456) to create a sample with children who were as comparable to the ECEAP sample of children as possible and had a wealth of information about their pretreatment characteristics. Households are defined as the members of the assistance unit. An assistance unit is comprised of anyone in the household who shares meals; this means that assistance unit members do not necessarily have a biological or familial relationship. However, the head of household's spouse and underage children are required to be in the same assistance unit if they live together, even if they do not share meals. If a child's other parent lives in the household, that individual must be included in the assistance unit as well.
Data received included (but was not limited to) the following information for assistance unit members:

- Socio-demographic information,
- Monthly information on earned/unearned income,
- Services received from DSHS, and
- Quarterly employment hours and earnings (for parents only).

We also received the following household information:

- Household composition and
- Census tract of residency.

A list of all eligible children (children receiving DSHS's services that were age eligible for ECEAP) and their fellow child household members were sent to OFM/ERDC to identify children for which we needed ECEAP participation information and education outcomes. Using its P20W longitudinal data system, ERDC matched ECEAP child-and site-level data (from the ELMS database) to K-12 data (from OSPI's CEDARS database). This allowed us to track children's enrollment between ECEAP and the K-12 system. ERDC then sent us personally unidentifiable datasets that we linked together using anonymous IDs.

Data Processing

DSHS Data Processing

We received data for eligible children for each month of the ECEAP program year, in addition to 20 months prior to the start of ECEAP. For example, for the 2002 ECEAP program year (September 2001 to August 2002), we received data for each month, from January 2000 to August 2002. Children were included if their household received Basic Food benefits for at least 12 of the 32 months. Children were included if their relationship to the head of household fit in one of the following categories: natural or adopted child, child under legal guardianship, former step or natural child, stepchild, or co-parent's child. Grandchildren of the head of households and other relationships (less than 5% of the total eligible child population) were excluded because it was not possible to confirm and assume guardianship responsibilities for other types of heads of households and create parental covariate variables.

[38] Data sources are summarized in Exhibit A1.

Baseline household information was pulled from the August prior to the start of the ECEAP program year, or from the closest prior month to August. If a child appeared in two households in a single month, that month was skipped in favor of a prior single household month. All members of a household that share meals reported during that month were included when measuring the household size and all household members' income was used to calculate household income and percent poverty level. These household members under 18 were also used when creating an indicator for out-of-home placement in the past 12 months.

Head of household information along with spouse or co-parent data (for children and co-parent's children, respectively) were used to create parental covariate variables including parent gender, primary language, race, ethnicity, education, employment status, marital status, and teen parent status.[39] Census tract information was based on the head of household's address as of August prior to the ECEAP program year. We received the 2010 census tract geoid and merged poverty data from the 2000 census and 2009-2018 ACS 5-year estimates[40] using a geographic crosswalk for the 2000 data from NHGIS.[41]

There were several reasons a child or their household could be excluded during initial data cleaning and sample selection.

- If a person had problematic birthdate data and their age could not be calculated.
- If an identifying variable was somehow attached to two people.
- If a child did not qualify for Basic Food before the start of the ECEAP program year, they could not be included in that academic year because we lacked household information at baseline.
- If they did not have at least 12 months of Basic Food documented during the ECEAP year and/or 20 months prior.
- If the head of the household was not a parent, co-parent, or stepparent.
- If the assistance unit had no head of household or had multiple spouses in the same month, these monthly observations were also excluded.

These criteria accounted for the largest drop in our sample (664,456 to 364,469).

The second-largest drop in our sample was related to educational outcome availability. Of the 364,469 children with usable household data, 332,338 were uniquely identified in the ERDC data.

[39] Parents were defined as teen parents if they were under 18 years old when the ECEAP eligible child was born.
[40] Manson, S., Schroeder, J., Van Riper, D., Kugler, T., & Ruggles, S. *IPUMS National Historical Geographic Information System: Version 16.0 [dataset].* Minneapolis, MN: IPUMS. 2021.
[41] Schroeder, J.P. (2007). Target-density weighting interpolation and uncertainty evaluation for temporal analysis of census data. *Geographical Analysis, 39*(3), 311–335.

Exhibit A3 describes the differences between the cleaned RDA data that we were and were unable to match to data provided by ERDC. Children who had data in both data sets are referred to as the analysis sample. Children who only appeared in the cleaned RDA data are the excluded sample. The analysis sample includes children who live in census tracts with slightly higher poverty rates, although the household income is slightly higher. They are more likely to be Hispanic or multiracial and less likely to be White, Asian or Pacific Islander, or report their race as "other." They were more likely to report having a disability.[42] They are much more likely to receive subsidies.

The head of household is more likely to be a female, more likely to have less than a high school degree, and less likely to have attended at least some college. The head of the household is more likely to report working part-time. The head of the household is also less likely to be married or partnered. The second parent in the household is less likely to be present.[43]

[42] Disability is defined here as the children receiving DDA services during the eligibility determination period or disability was indicated in the basic food services application.
It is unclear if children in the analysis sample are more likely to be disabled or if children in the analysis sample are more likely to go through the process of getting disabilities identified during the eligibility period.
[43] This may be a direct result of the head of household being less likely to be married or partnered.

Exhibit A3

Comparison of Analysis Sample and Excluded RDA Sample

Variable	Analysis sample		Excluded RDA sample		
	Mean	SD	Mean	SD	p-value
Household income (% poverty level)	88.602	97.378	87.601	106.070	0.081
Census tract poverty rate	0.171	0.104	0.163	0.102	0.000
Monthly Basic Food benefits	368.557	214.132	400.204	236.916	0.000
Hispanic	0.316	0.465	0.252	0.434	0.000
White	0.378	0.485	0.426	0.495	0.000
Multiracial	0.173	0.378	0.118	0.322	0.000
Black	0.056	0.230	0.056	0.230	0.805
Asian or Pacific Islander	0.030	0.172	0.045	0.207	0.000
Native American or Alaskan Native	0.013	0.111	0.013	0.112	0.813
Other race	0.014	0.117	0.029	0.169	0.000
Female	0.488	0.500	0.494	0.500	0.043
Primary language – Spanish	0.127	0.333	0.109	0.312	0.000
Primary language – Other	0.044	0.204	0.044	0.205	0.818
Disability	0.049	0.217	0.016	0.127	0.000
Working Connections Child Care subsidy	0.355	0.479	0.160	0.367	0.000
Temporary Assistance for Needy Families (TANF)	0.362	0.481	0.298	0.457	0.000
Number in household	3.969	1.592	4.282	1.752	0.000
Number in household under age 5	1.505	0.651	1.581	0.691	0.000
Number in household under age 18	2.486	1.347	2.690	1.501	0.000
Number in household over 65	0.002	0.043	0.002	0.044	0.580
Head of household female	0.909	0.287	0.863	0.344	0.000
Head of household – Less than high school	0.308	0.462	0.266	0.442	0.000
Head of household – High school graduate or equivalent	0.510	0.500	0.499	0.500	0.000
Head of household – Some college or more	0.182	0.386	0.235	0.424	0.000
Head of household – Unemployed	0.530	0.499	0.658	0.474	0.000
Head of household – Part time 1-34 hours/week	0.318	0.466	0.236	0.425	0.000
Head of household – Full time ≥ 35 hours/week	0.152	0.359	0.106	0.308	0.000
Head of household – Married/living together	0.330	0.470	0.439	0.496	0.000
Head of household – Divorced/separated/widowed	0.190	0.392	0.174	0.379	0.000
Head of household – Under 18 when child born	0.042	0.202	0.029	0.169	0.000
Second parent is present in household	0.301	0.459	0.414	0.493	0.000
Sample size	332,338		32,131		

ECEAP Data Processing

We received data from ERDC for all children under 18-years-old in the dataset. ERDC combined this data with data on all children participating in ECEAP. The data came from two separate databases, the current ELMS database and the historical EMS dataset. We decided to code individuals as participating in ECEAP if they were enrolled at any time during the academic year. We ran sensitivity analyses where we restricted our sample to children who were enrolled in ECEAP for at least 180 days. This did not significantly impact our results.

ECEAP participants were included in the analysis if they had reliable RDA data. Exhibit A4 describes the differences between the analysis sample and ECEAP participants who either were not in the RDA data or did not have the requisite family information. The administrative ECEAP eligibility data may be less reliable than the administrative data from RDA, so we do not use it in our analysis.

These results show that children included in the sample are more likely to have participated in ECEAP for two years and as three-year-olds than ECEAP participants who do not have RDA data. We also find that they are more likely to report being Black or White and less likely to report being Hispanic. ECEAP participants included in the sample were less likely to be American Indian, Asian and slightly more likely to report being Pacific Islander, but the sample sizes were too small to report separately.

Exhibit A4

Comparison of Analysis Sample and Excluded ECEAP Sample

Variable	Analysis sample		Excluded ECEAP sample		
	Mean	SD	Mean	SD	p-value
Enrolled as three- and four-year-old	0.204	0.403	0.173	0.378	0.000
Enrolled as three-year-old only	0.120	0.325	0.112	0.315	0.000
Enrolled as four-year-old only	0.671	0.469	0.713	0.452	0.000
Hispanic	0.276	0.447	0.384	0.486	0.000
White	0.418	0.493	0.312	0.463	0.000
Multiracial	0.108	0.310	0.108	0.310	0.958
Black	0.095	0.293	0.075	0.263	0.000
Female	0.495	0.500	0.492	0.500	0.137
Primary language – English	0.756	0.429	0.596	0.491	0.000
Primary language – Spanish	0.191	0.393	0.337	0.476	0.000
Primary language – Other	0.053	0.225	0.067	0.251	0.000
Resides with one parent	0.492	0.500	0.352	0.478	0.00
Resides with two parents	0.460	0.498	0.594	0.491	0.000
Resides with other	0.046	0.210	0.051	0.221	0.000
Sample size	87,002		70,896		

Note:
The table above implies that fewer than 60% of children who participated in ECEAP were omitted from the sample. This is because some children were enrolled in ECEAP whose data could not be used in the comparison analysis due to data errors.

Exhibits A5-A7 outline the differences between the ECEAP participants and the comparison group on the child, parent, and household characteristics before matching and merging to the Washington public school data (after the third step in Exhibit 5). School-level characteristics depend on the outcome analyzed, so a comparison between treatment and comparison children is not available in general.

Exhibit A5

Child Level Descriptive Statistics for ECEAP Participants and Non-Participants

Variable	Non-ECEAP participants		ECEAP participants		
	Mean	SD	Mean	SD	p-value
Age at initial ECEAP eligibility	3.082	0.275	3.081	0.273	0.26
Number of years eligible for ECEAP	1.872	0.334	1.874	0.332	0.24
Hispanic	0.306	0.461	0.359	0.480	0.00
White	0.387	0.487	0.343	0.475	0.00
Multiracial	0.174	0.379	0.171	0.376	0.06
Black	0.055	0.227	0.063	0.243	0.00
Asian or Pacific Islander	0.032	0.175	0.026	0.159	0.00
Native American	0.014	0.116	0.008	0.091	0.00
Other race	0.014	0.116	0.014	0.119	0.24
Female	0.486	0.500	0.496	0.500	0.00
Primary language – English	0.836	0.370	0.800	0.400	0.00
Primary language – Spanish	0.118	0.322	0.167	0.373	0.00
Primary language – Other	0.046	0.210	0.033	0.179	0.00
Disability	0.052	0.222	0.039	0.193	0.00
Out of home placement in past year	0.01	0.11	0.02	0.12	0.00
Temporary Assistance for Needy Families (TANF)	0.359	0.480	0.378	0.485	0.00
Working Connections Child Care subsidy	0.36	0.48	0.33	0.47	0.00
Diversion cash assistance	0.026	0.159	0.025	0.156	0.15
Sample size	267,022		65,316		

Notes:
Includes historical and recent cohorts.
Disability is defined here as the children receiving DDA services during the eligibility determination period or disability was indicated in the Basic Food services application.

Exhibit A6

Parent/Head of Household Level Descriptive Statistics
for ECEAP Participants and Non-Participants

Variable	Non-ECEAP participants		ECEAP participants		
	Mean	SD	Mean	SD	p-value
Female	0.908	0.290	0.916	0.277	0.00
Less than high school	0.302	0.459	0.333	0.471	0.00
High school graduate or equivalent	0.513	0.500	0.498	0.500	0.00
Some college or more	0.186	0.389	0.169	0.375	0.00
Unemployed	0.523	0.499	0.559	0.497	0.00
Part time 1-34 hours/week	0.320	0.467	0.308	0.462	0.00
Full time ≥ 35 hours/week	0.157	0.364	0.133	0.340	0.00
Married/living together	0.330	0.470	0.328	0.469	0.26
Divorced/separated/widowed	0.191	0.393	0.183	0.387	0.00
Under 18 when child born	0.042	0.200	0.044	0.206	0.01
Second parent is present in household	0.301	0.459	0.299	0.458	0.19
Sample size	267,022		65,316		

Notes:
Includes historical and recent cohorts.
Employment data are from the unemployment insurance data. As such, it does not include certain types of employees who may not be eligible for unemployment benefits, and therefore, may not be reported.

Exhibit A7

Household Descriptive Statistics for ECEAP Participants and Non-Participants

Variable	Non-ECEAP participants		ECEAP participants		
	Mean	SD	Mean	SD	p-value
Household income (% poverty level)	89.75	100.54	83.92	83.03	0.00
Monthly Basic Food benefits	366.43	215.46	377.25	208.38	0.00
Number in household	3.97	1.62	3.97	1.48	0.72
Number in household under age 5	1.51	0.65	1.50	0.65	0.13
Number in household under age 18	2.49	1.37	2.48	1.25	0.02
Number in household over 65	0.002	0.043	0.002	0.043	0.98
Any child in a household that received out of home placement services during the eligibility determination period	0.015	0.121	0.017	0.130	0.00
Census tract poverty rate	0.170	0.103	0.177	0.105	0.00
Sample size	267,022		65,316		

Note:
Includes historical and recent cohorts.

Outcome Processing and Construction

RDA Outcome Processing

Convictions. As noted in the main text of the report, we received information on youth convictions from AOC through RDA. We decided to use conviction instead of another metric of youth behavior (e.g., arrests or incarcerations) because it is the measure of guilt by a court of law. The use of arrests could potentially over-estimate crime and incarceration could under-estimate crime.

We separately report on misdemeanor convictions, felony convictions, and any convictions.

We begin counting convictions in the academic year a child turns 12[44] up through the end of the academic year a child turned 14, 15, 16, or 17. Ultimately, we treat this outcome as a binary variable rather than a count of convictions; this means the outcome can be interpreted as "at least one conviction by the end of the academic year they turn 17 years of age."

We also restrict the sample to only include children who were enrolled in a Washington high school in 9th grade to avoid the inclusion of children who moved out of the state prior to entering high school.

Exhibit A8 shows which cohorts are included in our analysis for each age. For example, our analysis on crime by 14 consists of all children born between September 1, 1996, and August 31, 2004.

[44] We used 12 as a cutoff because children are unlikely to be tried for committing a crime before they are 12. See RCW 9A.04.050.

Exhibit A8

Cohorts Included in Conviction by Age Analysis

		Age (by end of the academic year)																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		Grade (by end of the academic year)																	
		Born by Aug 31…				ECEAP (3)	ECEAP (4)	K	1	2	3	4	5	6	7	8	9	10	11	12
2014 study cohort	Historical cohort	**97**	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
		98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
		99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
		02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	
		03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19		
		04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19			
	Recent cohort	**05**	06	07	08	09	10	11	12	13	14	15	16	17	18	19				
		06	07	08	09	10	11	12	13	14	15	16	17	18	19					
		07	08	09	10	11	12	13	14	15	16	17	18	19						
		08	09	10	11	12	13	14	15	16	17	18	19							
		09	10	11	12	13	14	15	16	17	18	19								
		10	11	12	13	14	15	16	17	18	19									
		11	12	13	14	15	16	17	18	19										
		12	13	14	15	16	17	18	19											
		13	14	15	16	17	18	19												
		14	15	16	17	18	19													

Note:
The table shows the last two digits of the academic year.

We exclude conviction data from 2019 because there were numerous changes to the dataset which made this year less reliable. The most notable change was that King County no longer appear in our dataset.

Family Outcomes. We also received information on CPS involvement, parental wages, and parental hours worked from RDA.

Parental Wages and Hours Worked. All parent employment outcomes are measured by the increase or decrease relative to the year prior for the head of household, usually driven by changes in the child's mother's employment outcomes. For hours worked and wages earned we calculated the difference between the ECEAP academic year and the previous academic year. Our data was only available at the quarterly level rather than the monthly level, so there is a slight difference in months covered with this measure. An academic year so far has been defined as Sept 1st through August 31st. For our employment outcomes, it is defined as October 1st through September 30th.

All wage and employment data are from the unemployment insurance data. As such, it does not include certain types of employees who may not be eligible for unemployment benefits, and therefore, may not be reported. This includes independent contractors, self-employed individuals, and other exempt positions.[45] This may result in undercounting parental earnings and wages in our model. However, we believe that this will not result in much distortion in our results since we are examining the difference in earning and employment compared to the pre-period. As such, it would only cause a problem if parents were systematically switching between positions included and excluded from the unemployment insurance data as a result of participation in ECEAP.

CPS Involvement. A household is defined as having a CPS involvement if the assistance unit had any CPS-investigation or CPS-Family Assessment Response (FAR)[46] during a potential ECEAP enrollment period (the event did not have to involve the ECEAP eligible child). A positive estimate can be interpreted as a percentage point increase in CPS involvement. It is important to note that this is measured at the assistance unit level, rather than the child level. This is because CPS involvement in the case of any child could ultimately impact all children in the household.

ERDC/OSPI Outcome Processing

Test Scores. We calculated test scores based on the final state-mandated, standardized tests offered in the spring at the primary school attended. Test scores were only included when the testing grade matched the reporting grade for the student taking the test.

For children in the historical cohort group, test scores were based on their performance on the Washington Assessment of Student Learning (WASL) or Measurements of Student Progress (MSP) standardized test. Scores for children in the modern cohorts were based on their results from the Smarter Balanced Assessment (SBA) test. The general results do not include children who took the basic version of these tests. We ran an analysis including these children and the results were robust.

Assessments change over time and across grades; they cannot be directly compared in their raw form. To make them comparable within a cohort and across grades we converted test scores into z-scores. The z-score was calculated for each student by test grade, subject, and academic year as:

$$Z = (X - \mu)/\sigma$$

Where x=Is the student's test score
μ =the mean for the test for all students
σ = the standard deviation

This standardized test score for each test type, year, subject, and grade will have a mean score of zero and a standard deviation of one for the entire state of Washington. The mean and standard deviation of test scores in our sample do not equal 0 and 1, respectively because it is a subset of the population. If a student appeared in a grade multiple times, the test scores for the first time they appeared in the grade are reported.

[45] Employment Security Department. (2014). *Unemployment insurance tax information: A handbook for Washington state employers.* Olympia: WA. This is not an exhaustive list of exemptions. Additional information regarding the Unemployment Insurance program in Washington State is available from ESD.

[46] FAR is a CPS alternative response to a screened-in allegation of abuse or neglect. FAR focuses on children and youth safety along with the integrity and preservation of families when lower risk allegations of maltreatment have been screened-in for intervention.

High School Graduation. High school graduation was calculated in consultation with ERDC to closely match OSPI's definition.[47] To be included, students had to show up in the enrollment dataset at least once in 9th grade or higher. High school graduation captures whether the student graduated by their first expected graduation date (typically 4 years after 9th grade). To adjust for transfers in and out of high school, students with the following withdrawal codes were removed from our analysis:

- Deceased,
- Confirmed transfer to a medical facility with confirmation of educational services,
- Exited school due to medical reasons, is not receiving educational services, and
- Confirmed transfer out of Washington State.

[47] Weaver Randall, K., & Ireland, L. (2019). *Graduation and dropout statistics*. Office of Superintendent of Public Instruction. Olympia: WA.

IV. Methodology Summary

Matching on ECEAP Eligibility

For most of our analyses, we include children who participated in ECEAP when they were three years old, four years old, or both. To match them with suitable controls, we matched exactly on birth cohort and whether they were eligible when they were three years old only, four years old only, or both. To be consistent between cases and controls, we used the earliest year of baseline available. For most cases and controls, this was their three-year-old eligibility year information (see below).

Exhibit A9

Sample and Household Information Included in Academic and Behavioral Analysis

Age eligible for ECEAP	Age participated in ECEAP	Household information used for analysis
Three-years-old only	Did not participate	Year 3 eligibility determination information
	Three-years-old only	
Four-years-old only	Did not participate	Year 4 eligibility determination information
	Four-years-old only	
Both years	Did not participate	Year 3 eligibility determination information
	Three-years-old only	
	Four-years-old only	
	Three- and four-years-old	

Some of our outcomes occur during the year of ECEAP (parent employment). For this set of outcomes, we used information for children who participated in ECEAP at age four only. These participants and their comparison counterparts had to be eligible at age four. Children who were eligible in both years were only included if they did not participate in ECEAP at age three. This was done because the parental variables measure the difference in wages and hours worked. Children who participated in ECEAP when they were three would have been receiving treatment during the pre-period.

Exhibit A10

Sample and Household Information Included in Parental Outcomes Analysis

Age eligible for ECEAP	Age participated in ECEAP	Household information used for analysis
Four-years-old only	Did not participate	Year 4 eligibility determination information
	Four-years-old only	
Both years	Did not participate	Year 4 eligibility determination information
	Four-years-old only	

Matching Process

In addition to matching on the age attended ECEAP and the number of eligibility years, we further matched on more variables using Coarsened Exact Matching (CEM), and after matching we included more of these variables in regression models as controls. These variables are summarized in Exhibit A11 and A12.

Exhibit A11

Control Variables Used in Coarsened Exact Matching

Outcomes	Variables used in matching
All outcomes	Child characteristics
	• Birth cohort • Eligibility category (three-year-olds only, four-year-olds only, or both) • Sex • Binary race (White, BIPOC) • Disability status • Received childcare subsidy
	HOH/parent characteristics

	• Head of household education level
	  ○ Less than high school
	  ○ High school or GED
	  ○ Some college or more
	• Second parent present

	Household characteristics
	• Other children under age five present • Census tract, i.e., neighborhood poverty rate (limit to 4 bins)
WaKIDS	Geographic characteristics
	• Region (based on the residence when participating in ECEAP)
Academic outcomes (excluding WaKIDS)	School characteristics[#]
	• Test type (assessment results only) • Percent free and reduced-price meals (4 bins) • Percent White (4 bins) • School region
Behavioral outcomes	Geographic Characteristics
	• Juvenile justice court (based on the residence when participating in ECEAP)
Family outcomes	HOH/parent characteristics (parental employment and wages only)

	• HOH employment status[^] during the eligibility determination period
	  ○ Full-time
	  ○ Part-time
	  ○ Not in the labor force (NILF)/unemployed
	• Second parent employment status during the eligibility determination period
	  ○ Full-time
	  ○ Part-time
	  ○ Not in the labor force (NILF)/unemployed
	  ○ Second parent not present

	Geographic characteristics
	• Region (based on the residence when participating in ECEAP)

Notes:

[#] Not included in family outcome analyses.

[^] Based on average hours worked in the quarter prior to the first eligible ECEAP enrollment period. Full time = 35 hours per week or more, part-time =1 to 35 hours per week, and unemployed/not in labor force = 0 or no data.

All variables are based on their value during the eligibility determination period, not during ECEAP enrollment

Exhibit A12

Control Variables Used in Post Matching Regression

Outcome	Post-matching regression variable
All outcomes	Child characteristics

	• Birth cohort
	• Eligibility category (three-year-olds only, four-year-olds only, or both)
	• Sex
	• Race (White, Hispanic, multiracial, Black, Asian or Pacific Islander, other)
	• Disability status,
	• Primary language (English, Spanish, other)
	• Services received during the eligibility determination period:
	  ○ Childcare subsidy (flag)
	  ○ TANF (flag)
	  ○ DCA (flag)
	• Out of home placement (OOHP) during eligibility determination period (flag)

	HOH/parent characteristics

	• Head of household characteristics
	  ○ Female
	  ○ Marital status (what are categories)
	  ▪ Married or living together
	  ▪ Divorced, separated, or widowed
	  ▪ Never married, unknown, or missing
	  ○ Education level
	  ▪ Less than high school
	  ▪ High school or GED
	  ▪ Some college or More
	  ○ Employment status
	  ▪ Full-time
	  ▪ Part-time
	  ▪ Not in the labor force (NILF)/unemployed
	  ○ Teen parent status
	• Second parent present

	Household characteristics

	• Number of household members
	  ○ total,
	  ○ under 5,
	  ○ under 18, and
	  ○ over 65.
	• Any OOHP for members under 18
	• Monthly Basic Food benefits received (continuous)
	• Household % federal poverty level (continuous).

	Neighborhood characteristics
	• Poverty level

Notes:

^ Results were not sensitive to inclusion/exclusion of prior CPS involvement.

All variables are based on their value during the eligibility determination period, not during ECEAP enrollment.

Exhibit A12 (cont.)

Outcome	Post-matching regression variable
Academic outcomes	School characteristics
	• Test type (test scores only) • Percent Asian or pacific islander • Percent American Indian or Alaskan Native • Percent Black • Percent Hispanic • Percent White • Percent male • Percent FRPL
Behavioral outcomes	No additional controls
Parent outcomes	Parent characteristics
	• Total hours worked in the four quarters prior to ECEAP enrollment • Total wages in the four quarters prior to ECEAP enrollment
CPS involvement	Household characteristics
	• Any CPS involvement in the year prior to ECEAP enrollment^

Notes:
^Results were not sensitive to inclusion/exclusion of prior CPS involvement.
All variables are based on their value during the eligibility determination period, not during ECEAP enrollment.

School Outcomes Note

We excluded post-treatment controls when possible. As a result, the school controls used in the analysis included school information based on the school attended not during the outcome of interest but during the year before the student attended that school.

For Elementary School Outcomes. We controlled for the school characteristics of attended school in the year that the child would have been four-year-old (potentially enrolled in ECEAP).

For High School Outcomes. We controlled for the school characteristics of attending school in the year before 9th grade. For example, when evaluating outcomes for a 10th grader in the academic year 2017 we controlled for characteristics of the school from the academic year 2015, the year before the 10th grader started high school.

Behavioral Outcomes Note

Each age cut-off (14-, 15-, 16-, and 17-years-old) and conviction type (misdemeanor, felony, either) combination is a product of a separate analysis. For example, we ran a separate analysis for committing a misdemeanor as a 14-year-olds. For a list of variables matched on versus included in the models see Exhibits A11 and A12 above.

Family Outcomes Note

Each of the estimates for the different family outcomes and cohorts was products of separate analyses. For a list of variables matched on versus included in the models see Exhibits A11 and A12 above. School effects are not included in the models because outcome occurs before children enroll in elementary school.

V. Results

Exhibits A13-A20 show the results from the outcomes reported in the body of the report.

- Academic outcomes are reported in Exhibits A13-A17.
- Behavioral outcomes are reported in Exhibit A18.
- Contemporaneous family outcomes are reported in Exhibits A19-20.

Exhibit A13
WaKIDS Results

Grade	N	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
6 of 6 domains	79,983	0.2977	0.0357	0.0037	0.000
Socio-emotional	79,983	0.6739	0.0218	0.0037	0.000
Physical	79,983	0.7503	0.0327	0.0033	0.000
Language	79,983	0.6031	0.0329	0.0038	0.000
Cognitive	79,983	0.6023	0.0368	0.0039	0.000
Literacy	79,983	0.6368	0.0624	0.0037	0.000
Math	79,983	0.4865	0.0449	0.004	0.000

Exhibit A14

Test Scores (Z-Scores)

Historical cohort					
Grade	N	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
Math					
3	29,133	-0.458	0.026	0.011	0.023
4	27,664	-0.450	0.008	0.012	0.514
5	27,543	-0.449	-0.004	0.011	0.732
Reading					
3	28,989	-0.442	0.017	0.012	0.162
4	27,547	-0.410	0.004	0.012	0.735
5	22,152	-0.430	-0.004	0.013	0.771
Writing					
4	26,805	-0.374	-0.015	0.013	0.259

Recent cohort					
Grade	N	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
Math					
3	73,295	-0.3949	-0.016	0.008	0.038
4	61,800	-0.4011	-0.024	0.008	0.003
5	46,117	-0.4091	-0.026	0.009	0.006
ELA					
3	69,580	-0.4015	-0.026	0.008	0.001
4	62,077	-0.4100	-0.022	0.008	0.007
5	46,128	-0.4067	-0.027	0.009	0.004

Note:

^Individuals who passed based on the Basic score are not counted as meeting standards because their scores did not meet the statewide pass rate.

Exhibit A15

Impact of ECEAP Special Education Participation

Historical cohort					
Grade	N	Rate of the Comparison Group	Change related to ECEAP	Standard Error	p-value
K	26,305	0.150	-0.015	0.005	0.001
1	30,619	0.172	-0.014	0.005	0.002
2	32,843	0.197	-0.010	0.005	0.038
3	32,637	0.211	-0.004	0.005	0.408
4	31,607	0.218	0.000	0.005	0.985
5	30,836	0.225	-0.005	0.005	0.356
6	31,060	0.222	-0.010	0.005	0.043
7	30,738	0.219	-0.011	0.005	0.028
8	29,790	0.218	-0.009	0.005	0.073
9	24,914	0.217	-0.006	0.006	0.253
10	22,387	0.220	-0.003	0.006	0.643
11	16,159	0.215	-0.002	0.007	0.739
12	10,654	0.202	0.001	0.008	0.864

Recent cohort					
Grade	N	Rate of the Comparison Group	Change related to ECEAP	Standard Error	p-value
K	107,720	0.136	-0.004	0.002	0.058
1	107,628	0.150	-0.006	0.002	0.017
2	91,317	0.169	-0.009	0.003	0.001
3	74,938	0.187	-0.012	0.003	0.000
4	58,629	0.196	-0.009	0.004	0.015
5	43,124	0.199	-0.009	0.004	0.045
6	16,834	0.214	-0.012	0.007	0.093
7	15,462	0.200	-0.007	0.007	0.308
8	6,151	0.206	-0.014	0.011	0.194

Exhibit A16

High School Graduation—Historical Cohort Only

N	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
14,284	0.669	0.006	0.008	0.448

Exhibit A17

Juvenile Justice Convictions

Convictions through age	N	Rate of comparison group	Change related to ECEAP	Standard error	p-value
Misdemeanor convictions					
14	50,553	0.042	0.003	0.002	0.119
15	41,870	0.067	0.004	0.003	0.165
16	34,104	0.092	0.003	0.004	0.381
17	26,651	0.121	0.004	0.005	0.330
Felony convictions					
14	50,553	0.010	0.001	0.001	0.249
15	41,870	0.021	0.001	0.002	0.568
16	34,104	0.034	0.002	0.002	0.473
17	26,651	0.047	-0.001	0.003	0.722
Any convictions					
14	50,553	0.044	0.004	0.002	0.054
15	41,870	0.073	0.004	0.003	0.186
16	34,104	0.099	0.002	0.004	0.514
17	26,651	0.130	0.004	0.005	0.454

Exhibit A18

Parental Employment

Historical cohort				
Outcome	N	Coefficient	Standard error	p-value
Hours	32,303	-10.106	6.653	0.129
Wages	32,303	-194.481	76.729	0.011

Recent cohort				
Outcome	N	Coefficient	Standard error	p-value
Hours	111,928	2.277	3.927	0.562
Wages	111,928	-164.454	57.166	0.004

Exhibit A19

CPS involvement

Historical cohort				
N	Rate of comparison group	Margins	Standard error	p-value
53,351	0.124	0.007	0.003	0.039

Recent cohort				
N	Rate of comparison Group	Margins	Standard error	p-value
166,021	0.101	0.011	0.002	0.000

Sensitivity Analysis

Exhibits A20 – A34 show results from selected sensitivity analyses. In each case—

- CEM population, fully adjusted: Results reported in the body of the report (our preferred method).
- Unmatched population, unadjusted: Regression does not include any control variables.
- Unmatched population, fully adjusted: OLS regression with the same control variables used in the preferred method.
- CEM population, unadjusted: Results matched, but second stage regression does not include any control variables.
- CEM population, fully adjusted + distance: Adds distance between ECEAP center and residence as a control variable.
- CEM population, fully adjusted + restricted to 180 + days of ECEAP: Restricts sample to only children who were enrolled in ECEAP for at least 180 days.
- CEM population, fully adjusted + Head Start saturation: Adds a measure of Head Start saturation (measured by the number of Head Start slots, divided by the number of children below the federal poverty level under age five) at the CCA regional level (calculated for historical cohort group only).

Results from the sensitivity analysis are in the same order as the main results:

- Academic outcomes are reported in Exhibits A20-A27.
- Behavioral outcomes are reported in Exhibit A28-A30.
- Contemporaneous family outcomes are reported in Exhibits A31-A34.

We largely find that our results are robust to alternative model specifications. In most cases, alternative model specifications do not fundamentally change the magnitude or the statistical significance of our results.[48]

There are only a handful of notable exceptions. These are all related to one of the five sensitivity analyses we ran, the restriction of the sample to only children who were enrolled in ECEAP for at least 180 days.

- Academic outcomes
    - Test scores
        - The positive impact of ECEAP participation on 3rd-grade reading scores becomes statistically significant (the magnitude of the effect is similar) for the historical cohort group.
        - The negative impact of ECEAP participation on 3rd through 5th-grade test scores decreases in both statistical significance and magnitude for some outcomes for the recent cohort group.
            - The negative relationship between ECEAP participation and math scores is no longer statistically significant.
            - The negative relationship between ECEAP participation and ELA scores is less statistically significant.
- Contemporaneous family
    - Parental hours worked
        - The positive relationship between ECEAP participation and parental hours worked is now statistically significant for the recent cohort group (although it is still practically insignificant).
    - Parental wages
        - The negative relationship between ECEAP participation and parental wages for the historical and recent cohort groups is no longer statistically significant (in the original analysis, they were statistically significant but not practically significant).
    - CPS involvement
        - The positive relationship between ECEAP participation and CPS involvement is no longer statistically significant for the historical cohort group.

[48] Results do not change when controlling for Head Start saturation. This may be part because saturation is measured at too broad a geographic level, and thus may not measure if children truly have access to alternative care options.

Exhibit A20

Sensitivity Analysis: WaKIDS—6 of 6 Domains

Model description	N	ECEAP coefficient	Standard error	p-value	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
CEM population, fully adjusted	**79,983**	**0.173**	**0.018**	**0.000**	**0.298**	**0.036**	**0.004**	**0.000**
Unmatched population, unadjusted	89,741	0.109	0.017	0.000	0.303	0.024	0.004	0.000
Unmatched population, fully adjusted	89,735	0.166	0.017	0.000	0.300	0.034	0.004	0.000
CEM population, unadjusted	79,983	0.133	0.017	0.000	0.299	0.029	0.004	0.000
CEM population, fully adjusted + distance	79,994	0.173	0.018	0.000	0.298	0.036	0.004	0.000
CEM population, fully adjusted + restricted to 180+ days of ECEAP	78,249	0.203	0.018	0.000	0.298	0.042	0.004	0.000

Note:
The preferred model is **bolded**.

Exhibit A21

Sensitivity Analysis: 3rd Grade (Z-Scores)—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value
Math				
CEM population, fully adjusted	**29,133**	**0.026**	**0.012**	**0.023**
Unmatched population, unadjusted	64,449	-0.014	0.009	0.136
Unmatched population, fully adjusted	64,448	0.006	0.009	0.525
CEM population, unadjusted	29,133	0.020	0.012	0.097
CEM population, fully adjusted + distance	29,133	0.028	0.012	0.016
CEM population, fully adjusted + restricted to 180+ days of ECEAP	25,004	0.039	0.013	0.002
CEM population, fully adjusted + Head Start saturation	29,133	0.026	0.012	0.024
Reading				
CEM population, fully adjusted	**28,989**	**0.017**	**0.012**	**0.162**
Unmatched population, unadjusted	64,104	-0.005	0.010	0.609
Unmatched population, fully adjusted	64,103	0.006	0.010	0.565
CEM population, unadjusted	28,989	0.013	0.012	0.285
CEM population, fully adjusted + distance	28,989	0.019	0.012	0.116
CEM population, fully adjusted + restricted to 180+ days of ECEAP	24,908	0.023	0.013	0.075
CEM population, fully adjusted + Head Start saturation	28,989	0.016	0.012	0.169

Note:
The preferred model is **bolded**.

Exhibit A22

Sensitivity Analysis: 4th Grade (Z-Scores)—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value
Math				
CEM population, fully adjusted	**27,664**	**0.008**	**0.012**	**0.514**
Unmatched population, unadjusted	64,232	-0.021	0.009	0.020
Unmatched population, fully adjusted	64,231	0.002	0.009	0.868
CEM population, unadjusted	27,664	0.001	0.012	0.967
CEM population, fully adjusted + distance	27,664	0.009	0.012	0.447
CEM population, fully adjusted + restricted to 180+ days of ECEAP	23,544	0.009	0.013	0.483
CEM population, fully adjusted + Head Start saturation	27,664	0.008	0.012	0.506
Reading				
CEM population, fully adjusted	**27,547**	**0.004**	**0.012**	**0.735**
Unmatched population, unadjusted	63,072	-0.012	0.010	0.211
Unmatched population, fully adjusted	63,071	0.003	0.009	0.740
CEM population, unadjusted	27,547	-0.002	0.012	0.896
CEM population, fully adjusted + distance	27,547	0.007	0.012	0.578
CEM population, fully adjusted + restricted to 180+ days of ECEAP	23,399	0.002	0.013	0.890
CEM population, fully adjusted + Head Start saturation	27,547	0.004	0.012	0.744
Writing				
CEM population, fully adjusted	**26,805**	**-0.015**	**0.013**	**0.259**
Unmatched population, unadjusted	61,825	-0.028	0.011	0.008
Unmatched population, fully adjusted	61,824	-0.013	0.010	0.208
CEM population, unadjusted	26,805	-0.024	0.014	0.076
CEM population, fully adjusted + distance	26,805	-0.014	0.013	0.300
CEM population, fully adjusted + restricted to 180+ days of ECEAP	22,747	0.000	0.015	0.983
CEM population, fully adjusted + Head Start saturation	26,805	-0.014	0.012	0.279

Note:
The preferred model is **bolded**.

Exhibit A23

Sensitivity Analysis: 5th Grade (Z-Scores)—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value
Math				
CEM population, fully adjusted	**27,543**	**-0.004**	**0.011**	**0.732**
Unmatched population, unadjusted	63,648	-0.025	0.009	0.007
Unmatched population, fully adjusted	63,647	-0.002	0.009	0.796
CEM population, unadjusted	27,543	-0.013	0.012	0.284
CEM population, fully adjusted + distance	27,543	-0.003	0.011	0.783
CEM population, fully adjusted + restricted to 180+ days of ECEAP	23,127	0.005	0.013	0.692
CEM population, fully adjusted + Head Start saturation	27,543	-0.004	0.011	0.701
Reading				
CEM population, fully adjusted	**22,152**	**-0.004**	**0.013**	**0.781**
Unmatched population, unadjusted	51,773	-0.023	0.011	0.033
Unmatched population, fully adjusted	51,772	-0.007	0.010	0.529
CEM population, unadjusted	22,152	-0.010	0.014	0.449
CEM population, fully adjusted + distance	22,152	-0.003	0.013	0.815
CEM population, fully adjusted + restricted to 180+ days of ECEAP	18,457	0.004	0.015	0.794
CEM population, fully adjusted + Head Start saturation	22,152	-0.004	0.0134	0.752

Note:
The preferred model is **bolded**.

Exhibit A24

Sensitivity Analysis: 3rd Grade (Z-Scores)—Recent Cohort

Model description	N	ECEAP coefficient	Standard error	p-value
Math				
CEM population, fully adjusted	**73,295**	**-0.016**	**0.008**	**0.038**
Unmatched population, unadjusted	111,561	-0.062	0.007	0.000
Unmatched population, fully adjusted	111,558	-0.020	0.007	0.004
CEM population, unadjusted	73,295	-0.031	0.008	0.000
CEM population, fully adjusted + distance	73,295	-0.015	0.008	0.048
CEM population, fully adjusted + restricted to 180+ days of ECEAP	69,437	-0.007	0.008	0.389
ELA				
CEM population, fully adjusted	**69,580**	**-0.026**	**0.008**	**0.001**
Unmatched population, unadjusted	103,841	-0.073	0.007	0.000
Unmatched population, fully adjusted	103,838	-0.030	0.007	0.000
CEM population, unadjusted	69,580	-0.041	0.008	0.000
CEM population, fully adjusted + distance	69,580	-0.026	0.008	0.001
CEM population, fully adjusted + restricted to 180+ days of ECEAP	66,112	-0.020	0.008	0.014

Note:
The preferred model is **bolded**.

Exhibit A25

Sensitivity Analysis: 4th Grade (Z-Scores)—Recent Cohort

Model description	N	ECEAP coefficient	Standard error	p-value
Math				
CEM population, fully adjusted	**61,800**	**-0.024**	**0.008**	**0.003**
Unmatched population, unadjusted	93,955	-0.063	0.008	0.000
Unmatched population, fully adjusted	93,954	-0.020	0.007	0.008
CEM population, unadjusted	61,800	-0.037	0.009	0.000
CEM population, fully adjusted + distance	61,800	-0.025	0.008	0.003
CEM population, fully adjusted + restricted to 180+ days of ECEAP	57,977	-0.014	0.009	0.099
ELA				
CEM population, fully adjusted	**62,077**	**-0.022**	**0.008**	**0.007**
Unmatched population, unadjusted	94,160	-0.064	0.008	0.000
Unmatched population, fully adjusted	94,159	-0.022	0.007	0.003
CEM population, unadjusted	62,077	-0.035	0.009	0.000
CEM population, fully adjusted + distance	62,077	-0.023	0.008	0.004
CEM population, fully adjusted + restricted to 180+ days of ECEAP	58,191	-0.017	0.009	0.046

Note:
The preferred model is **bolded**.

Exhibit A26

Sensitivity Analysis: 5th Grade (Z-Scores)—Recent Cohort

Model description	N	ECEAP coefficient	Standard error	p-value
Math				
CEM population, fully adjusted	**46,117**	**-0.026**	**0.009**	**0.006**
Unmatched population, unadjusted	71,278	-0.064	0.009	0.000
Unmatched population, fully adjusted	71,277	-0.021	0.009	0.015
CEM population, unadjusted	46,117	-0.038	0.010	0.000
CEM population, fully adjusted + distance	46,117	-0.026	0.009	0.005
CEM population, fully adjusted + restricted to 180+ days of ECEAP	42,649	-0.017	0.010	0.098
ELA				
CEM population, fully adjusted	**46,128**	**-0.027**	**0.009**	**0.004**
Unmatched population, unadjusted	71,303	-0.063	0.009	0.000
Unmatched population, fully adjusted	71,302	-0.024	0.008	0.005
CEM population, unadjusted	46,128	-0.036	0.010	0.000
CEM population, fully adjusted + distance	46,128	-0.028	0.009	0.003
CEM population, fully adjusted + restricted to 180+ days of ECEAP	42,626	-0.020	0.010	0.046

Note:
The preferred model is **bolded**.

Exhibit A27

Sensitivity Analysis: High School Graduation—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
CEM population, fully adjusted	**14,284**	**0.031**	**0.040**	**0.449**	**0.669**	**0.006**	**0.008**	**0.448**
Unmatched population, unadjusted	34,394	0.003	0.030	0.917	0.672	0.001	0.007	0.917
Unmatched population, fully adjusted	34,389	0.026	0.031	0.402	0.671	0.005	0.006	0.401
CEM population, unadjusted	14,284	0.024	0.039	0.538	0.670	0.005	0.009	0.537
CEM population, fully adjusted + distance	14,284	0.035	0.040	0.381	0.669	0.007	0.008	0.380
CEM population, fully adjusted + restricted to 180+ days of ECEAP	11,869	0.046	0.045	0.314	0.668	0.010	0.009	0.313
CEM population, fully adjusted + Head Start saturation	14,284	0.038	0.040	0.343	0.671	0.008	0.008	0.342

Note:
The preferred model is **bolded**.

Exhibit A28

Sensitivity Analysis: 17-Year-Old Convictions (Misdemeanors)—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
CEM population, fully adjusted	**26,651**	**0.043**	**0.043**	**0.326**	**0.121**	**0.004**	**0.005**	**0.330**
Unmatched population, unadjusted	49,403	0.060	0.036	0.093	0.115	0.006	0.004	0.097
Unmatched population, fully adjusted	49,402	0.022	0.037	0.556	0.116	0.002	0.004	0.558
CEM population, unadjusted	26,651	0.041	0.042	0.339	0.121	0.004	0.005	0.342
CEM population, fully adjusted + distance	26,651	0.045	0.043	0.303	0.121	0.005	0.005	0.306
CEM population, fully adjusted + restricted to 180+ days of ECEAP	22,369	0.050	0.050	0.315	0.118	0.005	0.005	0.319
CEM population, fully adjusted + Head Start saturation	26,651	0.045	0.043	0.305	0.121	0.005	0.005	0.308

Note:
The preferred model is **bolded**.

Exhibit A29

Sensitivity Analysis: 17-Year-Old Convictions (Felonies)—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
CEM population, fully adjusted	**26,651**	**-0.024**	**0.068**	**0.723**	**0.047**	**-0.001**	**0.003**	**0.722**
Unmatched population, unadjusted	49,403	0.066	0.058	0.252	0.041	0.003	0.002	0.261
Unmatched population, fully adjusted	49,402	0.050	0.059	0.393	0.041	0.002	0.002	0.400
CEM population, unadjusted	26,651	-0.021	0.066	0.754	0.047	-0.001	0.003	0.753
CEM population, fully adjusted + distance	26,651	-0.024	0.068	0.720	0.047	-0.001	0.003	0.718
CEM population, fully adjusted + restricted to 180+ days of ECEAP	22,369	-0.024	0.078	0.758	0.046	-0.001	0.003	0.756
CEM population, fully adjusted + Head Start saturation	26,651	-0.024	0.068	0.726	0.047	-0.001	0.003	0.725

Note:
The preferred model is **bolded**.

Exhibit A30

Sensitivity Analysis: 17-Year-Old Convictions (Any)—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
CEM population, fully adjusted	**26,651**	**0.032**	**0.042**	**0.451**	**0.130**	**0.004**	**0.005**	**0.454**
Unmatched population, unadjusted	49,403	0.056	0.035	0.109	0.124	0.006	0.004	0.114
Unmatched population, fully adjusted	49,402	0.043	0.036	0.227	0.124	0.005	0.004	0.231
CEM population, unadjusted	26,651	0.030	0.041	0.463	0.130	0.003	0.005	0.466
CEM population, fully adjusted + distance	26,651	0.031	0.042	0.465	0.130	0.003	0.005	0.468
CEM population, fully adjusted + restricted to 180+ days of ECEAP	22,369	0.036	0.048	0.459	0.128	0.004	0.005	0.462
CEM population, fully adjusted + Head Start saturation	26,651	0.034	0.042	0.420	0.130	0.004	0.005	0.422

Note:
The preferred model is **bolded**.

Exhibit A31

Sensitivity Analysis: Parental Employment—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value
Change in hours worked				
CEM population, fully adjusted	**32,303**	**-10.11**	**6.65**	**0.129**
Unmatched population, unadjusted	69,958	-10.59	5.98	0.076
Unmatched population, fully adjusted	69,957	-8.34	5.67	0.142
CEM population, unadjusted	32,303	-5.11	6.97	0.463
CEM population, fully adjusted + distance	32,173	-10.43	6.67	0.118
CEM population, fully adjusted + restricted to 180+ days of ECEAP	26,846	8.01	7.47	0.284
CEM population, fully adjusted + Head Start saturation	32,303	-10.10	6.65	0.129
Change in wages earned				
CEM population, fully adjusted	**32,303**	**-194.48**	**76.73**	**0.011**
Unmatched population, unadjusted	69,958	-236.48	70.27	0.001
Unmatched population, fully adjusted	69,957	-172.98	67.97	0.011
CEM population, unadjusted	32,303	-151.06	79.19	0.056
CEM population, fully adjusted + distance	32,173	-200.94	76.97	0.009
CEM population, fully adjusted + restricted to 180+ days of ECEAP	26,846	-13.49	87.47	0.877
CEM population, fully adjusted + Head Start saturation	32,303	-196.65	76.73	0.010

Note:

The preferred model is **bolded**.

Exhibit A32

Sensitivity Analysis: Parental Employment—Recent Cohort)

Model description	N	ECEAP coefficient	Standard error	p-value
Change in hours worked				
CEM population, fully adjusted	**111,928**	**2.28**	**3.93**	**0.562**
Unmatched population, unadjusted	174,014	-1.95	3.86	0.614
Unmatched population, fully adjusted	174,007	0.88	3.66	0.810
CEM population, unadjusted	111,928	6.41	4.15	0.123
CEM population, fully adjusted + distance	111,774	2.15	3.93	0.585
CEM population, fully adjusted + restricted to 180+ days of ECEAP	105,829	8.75	4.14	0.035
Change in wages earned				
CEM population, fully adjusted	**111,928**	**-164.45**	**57.17**	**0.004**
Unmatched population, unadjusted	174,014	-183.31	56.06	0.001
Unmatched population, fully adjusted	174,007	-202.33	54.17	0.000
CEM population, unadjusted	111,928	-62.99	59.21	0.287
CEM population, fully adjusted + distance	111,774	-165.76	57.22	0.004
CEM population, fully adjusted + restricted to 180+ days of ECEAP	105,829	-93.49	60.70	0.124

Note:

The preferred model is **bolded**.

Exhibit A33

Sensitivity Analysis: CPS Involvement—Historical Cohort

Model description	N	ECEAP coefficient	Standard error	p-value	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
CEM population, fully adjusted	**53351**	**0.076**	**0.036**	**0.036**	**0.124**	**0.007**	**0.003**	**0.039**
Unmatched population, unadjusted	75394	0.046	0.033	0.165	0.126	0.005	0.003	0.169
Unmatched population, fully adjusted	75393	0.063	0.034	0.063	0.125	0.006	0.003	0.066
CEM population, unadjusted	53351	0.063	0.036	0.075	0.124	0.006	0.004	0.078
CEM population, fully adjusted + distance	53189	0.081	0.036	0.027	0.123	0.008	0.003	0.029
CEM population, fully adjusted + restricted to 180+ days of ECEAP	46880	0.058	0.042	0.166	0.120	0.005	0.004	0.171
CEM population, fully adjusted + Head Start saturation	53351	0.079	0.036	0.029	0.124	0.007	0.003	0.031

Exhibit A34

Sensitivity Analysis: CPS Involvement—Recent Cohort

Model description	N	ECEAP coefficient	Standard error	p-value	Rate of the comparison group	Change related to ECEAP	Standard error	p-value
CEM population, fully adjusted	**166021**	**0.135**	**0.022**	**0.000**	**0.101**	**0.011**	**0.002**	**0.000**
Unmatched population, unadjusted	192427	0.103	0.021	0.000	0.105	0.009	0.002	0.000
Unmatched population, fully adjusted	192419	0.121	0.021	0.000	0.105	0.010	0.002	0.000
CEM population, unadjusted	166021	0.136	0.022	0.000	0.101	0.012	0.002	0.000
CEM population, fully adjusted + distance	165845	0.140	0.022	0.000	0.101	0.011	0.002	0.000
CEM population, full adjusted + restricted to 180+ days of ECEAP	160109	0.099	0.024	0.000	0.100	0.008	0.002	0.000

VI. Discussion of Methodological Changes

The main concern is that the results found in this report are inconsistent with those found in the 2014 evaluation of ECEAP published by WSIPP in 2014. The 2014 report found that ECEAP participation was positively related to a statistically significant increase in 3rd-, 4th-, and 5th-grade assessment scores. In addition, we found that this positive relationship was persistent, and did not significantly decay over the three grades analyzed. The current study finds that ECEAP does not have a statistically significant impact on test scores for the historical cohorts. We also find that ECEAP has a small negative, but statistically significant, association with test scores for the recent cohorts.

For the current report, we made several changes made to the cohort construction, control variables used, and final methodology. This section explores the history behind the choice to change the methodology for the 2022 report, the rationale for the changes in the study design, and the main differences between the two reports. We believe that the difference in results stems primarily from the change in methodology and discuss why the updated methodology is preferred.

The 2014 report relied on an Instrumental Variable approach. This report uses CEM to estimate the relationship between ECEAP participation and later outcomes.

Why Did We Choose to Change our Methodology?

We decided to change our methodology because the magnitude and persistence of the results in the 2014 report were out of step with results produced by other researchers studying the impact of ECE programs. This was acknowledged in the original report and further corroborated by results in the 2019 meta-analysis. We also believed that there were some limitations in the Instrumental Variable research design used in the report.

Having results that are different from the rest of the literature does not mean that the original authors did anything wrong, and indeed we believe that the results from the report we produced were the best that we could do given the data available to us at the time. However, with the new report, we had the opportunity to collect additional data which we believed would help us to better understand the relationship between ECEAP and later outcomes than the methodology used in the original report.

Unfortunately, the data did not support methodologies that would allow us to determine the causal relationship between ECEAP and later outcomes. We used CEM as our best alternative. We hoped that the improved methodology would act as a robustness check to the original results.

Original Methodology and Critique

The 2014 study used the geographic distance from each ECEAP eligible child's residence to the nearest ECEAP center as the instrumental variable to estimate ECEAP participation. We conducted a maximum likelihood estimation of a two-stage instrumental variable model.

1) In the first stage equation, we estimated the probability of ECEAP participation using distance from each child's home when they were three or four years old to the nearest ECEAP provider[49] as an instrumental variable.[50]
2) In the second stage equation, we modeled achievement test scores including the estimated probability of ECEAP participation generated from the first stage equation.[51]

The exact distance measure was constructed (as the crow flies) by using the distance between the midpoint of the census tract for each household home address and the latitude and longitude coordinates of the closest ECEAP provider.[52]

The rationale for using an Instrumental Variable approach is that the statistical strategy reduces bias from unmeasured variables if the following conditions are met:

1) The distance from an ECEAP center (the instrument) has a causal effect on ECEAP enrollment (the treatment).
2) The distance from an ECEAP center affects test scores (the outcome variable) only through ECEAP attendance.
    a. The distance between residence and ECEAP centers does not have a direct influence on test scores, and
    b. No variables which impact both the distance from an ECEAP center and test scores are not controlled for in our analysis.

Another consideration related to (2b) is that if the instrument (distance from ECEAP center) only impacts a subset of the population, then the results measured by the instrument will be internally valid (accurately capture the relationship between ECEAP and outcomes for the subpopulation) but may not generalize to the entire treated population.[53]

As stated previously, there are many unobserved factors (e.g., a parent's motivation to ensure that their child is academically prepared for kindergarten) which may impact ECEAP participation. The authors of the 2014 report believed that distance from the ECEAP center was likely to be correlated with a parent's choice to enroll a child in ECEAP but uncorrelated with unobserved characteristics (like parental motivation). At the time of publication, other researchers had successfully used distance as an instrument.[54]

---

[49] This was not necessarily the ECEAP center attended by ECEAP participants.
[50] In addition to the instrument, the equation includes a set of covariates drawn from the DSHS Basic Food database which were temporally aligned with the ECEAP enrollment decision. The equation included characteristics of the child such as gender, race, ethnicity disability status, language spoken, and relationship to the head of household. Additional covariates include household level characteristics such as income, type of Basic Food subsidy program, childcare subsidy status, homelessness status, household size and age composition, and neighborhood poverty rate.
[51] In addition to the covariates included in the first stage, we include several other covariates. For each student, we include binary variables indicating whether they received free or reduced-price meal and disability status, a variable indicating percent of the academic year the child attended and the total number of schools attended during the current academic year. Finally, we included school-level fixed effects and birth cohort fixed effects.
[52] A technical problem with using this distance data is that we do not have a direct measure of the child's residence. The accuracy of the distance measure, and therefore its ability to predict ECEAP enrollment, will be inversely related to the size of the census tract. The size of the census tract is correlated with many other factors not controlled for in our analysis, such as urbanicity, and may be correlated with other factors which could impact academic outcomes. In addition, using distance as the crow flies is that it may not be functionally meaningful because it does not capture the actual time it might take for parents to go to the center (instead of measuring driving time, public transportation time etc.).
[53] Pizer, S.D. (2016). Falsification testing of instrumental variables methods for comparative effectiveness research. *Health Services Research, 51*(2), 790-811.
[54] Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling, in: N. Louis, E. Christofides, K. Grant, & R. Swidinsky, (Eds.), *Aspects of labour market behaviour: Essays in honour of John Vanderkamp* (University of Toronto Press, Toronto, Canada) pp. 201-222.

Theoretical Problems with Distance as an Instrument

One concern that we had when revising our analysis was that distance may not have a strong enough impact on ECEAP enrollment (violates condition one). The other concern was that there was a high likelihood that other omitted factors were correlated with both the distance from an ECEAP center and related to subsequent academic and behavioral outcomes (violates condition two). ECEAP centers are not randomly distributed throughout a community. This is by design, as DCYF and ECEAP target outreach and expansion to communities of greatest need. Such targeting further complicates sample construction. It is probable that there may be fundamental differences between communities that have access to ECEAP centers and those that do not. These community characteristics may also impact later outcomes, especially if the children reside in a nearby community throughout their education. It is also probable that individuals who are induced to participate in ECEAP because of the center's proximity to their house are fundamentally different from other ECEAP participants and so the results will not generalize to the entire ECEAP population.

These potential violations of the assumptions of this methodology mean that the results from this methodology may be biased. Therefore, for this report, we elected to use a different methodology that did not suffer from these limitations.

Updated Methodology and Limitations

As stated in the body of the report, we used CEM to match ECEAP children to comparable non-participants. ECEAP participants are then matched to children who share the same characteristics, across all variables used in the matching (they will have the same age, race, parental years of education, etc.) Children who do not have matches are dropped from the sample.

After matching, we run a regression modeling the effect of ECEAP controlling for the variables included in the CEM, additional child, parent, and household characteristics on which we were unable to match, and school characteristics[55] (when applicable)[56] while adjusting for weights generated by the coarsened exact matching. We used linear regression for continuous outcomes (e.g., math assessment), and a logistic regression for binary outcomes (e.g., high school graduation).

What are the Limitations

CEM can produce poor results in the following scenarios:

1) Variable Coarsening is too broad,
2) Variable Coarsening to too narrow, and
3) Key variables are omitted.

If the variable bins are too broad, critical variation within the category may be missed. Individuals in the treatment group may then be matched to individuals in the comparison group who are not very similar to them. This can be partially fixed in the second stage regression.

If the bins are too narrow, then it increases the likelihood that some individuals will not be matched at all.

[55] We control for the following school characteristics: Percent Asian, percent Native American or Alaskan Native, percent Black or African American, percent Hispanic, percent White, percent more than one race, percent male, and percent receiving free or reduced-price lunch.

[56] We do not control for school characteristics when analyzing family outcomes.

The third limitation is a limitation of all matching techniques. We believe that it is highly likely to be of concern in our current analysis.

We attempted to balance bin size so that they were not so broad as to be useful but also not so narrow as to throw out large portions of our sample.

The remaining concern that we have about this methodology is the potential for omitted variables. These variables, which should be included but cannot be due to data limitations, prevent us from making causal claims. These concerns are discussed in detail throughout the body of the report.[57]

Other Model Difference Between the 2014 and Current Evaluation

Exhibit A35 highlights the main differences in samples in the two studies.

Exhibit A35

Overview of Differences in Sample Between 2014 and 2022 Report

	2014 study	Current study
Outcome measure	Test scores (WASL and MSP)	Test scores (WASL and MSP)
Treatment	Attended ECEAP as a three- or four-year-old	Attended ECEAP as a three- or four-year-old
Sample years	Children who were born between September **1999** and August 2004.	Children who were born between September **1996** and August 2004.
Eligibility	Received food stamp benefits for at least 12 months during a 30-month window beginning March 1 prior to August 31 of the year in which a child turns three years old.	Received food stamp benefits for at least 12 months over the ECEAP program year and prior 20 months.
Cohort construction	Children who **received Basic Food benefits** when they were **three or four years old** and subsequently attended Washington State public schools.	Children who **received DSHS services**^ during the **ECEAP eligibility determination period,#** **are living with at least one parent**, and subsequently attended Washington State public schools.

Notes:
Differences between the reports are **bolded.**

^ Changes to state and federal laws prevented RDA from sharing PII information to ERDC on Basic Food recipients for this project. We had to broaden the sample so that recipients of specific services could not be identified. We are able to make sure that our sample is relatively comparable to the original because we have information on household income and services received.

We use a base month of August prior to ECEAP enrollment because parent characteristics (such as employment and wages) would not be impacted by ECEAP participation. The original study used the first month of ECEAP participation as the default month. We expect to do additional sensitivity tests to see if switching months change our results.

First, there were slight differences in the eligibility determination window and the cohort construction. While the original study required children to receive Basic Food benefits for at least 12 months out of the 30-month window before they turned 3, we required children to be eligible for at least 12 months over the ECEAP program year and prior 20 months. We do not believe that the slight variations in the months used to determine ECEAP eligibility impacted the final outcomes. Some sensitivity analysis on the number of months in the eligibility window used to determine ECEAP eligibility showed that it did not meaningfully change our results.

[57] See Section VI—Summary and Limitations.

There were also changes made to the cohort construction across reports. Although we requested children who received more services from DSHS than just Basic Food, this decision was made to comply with changes to state and federal laws.[58] The final sample for both studies only included children who received Basic Food services. This was a trivial change and should not have impacted results.

The third change to our analysis was to restrict the sample to only include children who lived with at least one parent. We are unable to replicate the CEM method on the full sample because we do not have enough information on children not living with parents, however, we believe that this is not the primary reason for the difference in results between the reports, as less than 3% of household heads in the original study were not a parent.

Control Differences

Exhibits A36-A38 highlight the differences in variables used in the analysis in the two reports.

[58] Privacy laws prevented DSHS from sending PII needed for the child match to ERDC because ERDC would have a list of children receiving Basic Food services. A larger list of children receiving any services was sent so that ERDC would not be able to identify.

Exhibit A36

Overview of Differences in Variables Used in the 2014 and 2022 Report

2014 study	Current study
Distance from nearest ECEAP provider	Not included in main results
Log (net income per capita)	**Household FPL**^
Census tract poverty rate	Included
Race	**Race (coded differently)***
Hispanic	Included
Female	Female
Primary language	Primary language
Disability status	Disability status
Number in the household under 2	Number of children <5
Number in household 3 to 5	
Number in household 6 to 12	Number of children <18
Number in household 13 to 17	
Number in household 65 and over	Included
Number in household	Number of HH members (any)
The household head is a grandparent	**NA#**
The household head is not parent or grandparent	
Basic Food sub program C	**Type of program services received**^^ **We control whether the child received the following:** Temporary Assistance for Needy Families (TANF) Diversion Cash Assistance (DCA) Out-of-home placement (OOHP)
Basic Food sub program G	
Basic Food sub program other	
Working Connections Child Care (WCCC)	Included
School-level fixed effects	Included
Birth cohort fixed effects	Included

Notes:
Differences between the reports are **bolded**.
^The original report used log income per capita. We used household FPL, which adjusts household income based on the federal poverty level at the time of ECEAP eligibility determination.
*We use mutually exclusive race categories. The original study did not use mutually exclusive racial categories. However, we ran our analysis using the racial categories in the original study and it did not significantly impact our results.
#We restricted our sample to only include children living in a household where the head of household is either a parent or stepparent. As such, there are many fewer categories in the relationship to head of household category than are included in the 2014 study. We do not think that this will significantly impact our results because 97% of children in the original sample were living in a household headed by a parent.
^^We do have information on services received by households.

Exhibit A37

Variables Included in 2014 Study Only

2014 study	Current study
Homeless flag	**Not included in the analysis**
Disability status (at grade five)	
Grade five FRPL	
Percent of grade five academic year at testing school	
Number of schools attending during grade five	

Note:
Differences between the reports are **bolded**.

Exhibit A38

Variables Included in 2022 Study Only

2014 study	Current study
Parental characteristics **(Not included in the analysis)**	**Mother/HOH education (less than HS, GED/12th grade completed, some college or more)**
	Mother/HOH employment status quarter of enrollment (full employment, part-time, unemployed/missing)
	Mother/HOH marital status (married and LT, divorced/separated, never married/unknown/missing)
	Mother/HOH is a teen parent (when ECEAP eligible child was born)
	Father/second parent is present in HH
Household OOHP **(Not included in the analysis)**	**Anyone under 18 in HH receives any OOHP services in the previous months**

Note:
Differences between the reports are **bolded**.

Apart from homelessness, variables included in the 2014 study that were not controlled for in the 2022 study in at least a modified way were grade five characteristics. These variables were intentionally omitted from the 2022 analysis because they are post-treatment characteristics. The standard convention has moved away from controlling for characteristics that occur after treatment (in this case ECEAP participation) because it can bias the results.

As noted, this study includes controls for a variety of parental characteristics not previously controlled for. Education literature suggests that parental attributes may impact children's academic performance. We added these characteristics to our model to improve the match between the treatment and comparison groups.

Application of the IV Methodology to the 2022 Cohort Group

We wanted to check to make sure that differences observed in the analysis were caused by methodical changes and not other changes in the research design. To do this, roughly replicate the IV methodology from the 2014 study and apply it to the full historical cohort group from 2022.

We did not request the same variables, and therefore could not perfectly replicate the analysis.
The current report expanded the number of cohorts used in our analysis as well as the outcomes measured. When examining the difference in reports, we focus on the assessment scores for the historical cohort only. The historical cohort includes three additional years of students not included in the 2014 study. We do not restrict our analysis to the cohorts used in the 2014 study.

Replication

Exhibit A39 lists the controls included in our replication

Exhibit A39

Overview of Variables Used in Replication

Variable list
Distance from nearest ECEAP provider
Household FPL^
Census tract poverty rate
Race (coded differently) #
Hispanic
Female
Primary language
Disability status
Number of children <5
Number of children <18
Number of HH members (any)
Type of program services received* • TANF • DCA • OOHP
Childcare subsidy receipt (WCCC)
School-level fixed effects
Birth-cohort fixed effects

Notes:
Variables modified from the 2014 report but controlling for the same characteristic, are **bolded**.
^ The original report used log income per capita. We used household FPL, which adjusts household income based on the federal poverty level at the time of ECEAP eligibility determination.
# We use mutually exclusive race categories. The original study did not use mutually exclusive racial categories. However, we ran our analysis using the racial categories in the original study and it did not significantly impact our results.
* We did not receive information on the type of food benefits received. We do have information on services received by households.

Results. Exhibit A40 summarizes the results from our replication.[59] We were able to partially replicate the results from the 2014 results using the current data. We generally replicate the finding that ECEAP has a positive and statistically significant impact on test scores, but there are some differences. The most obvious of which is that the coefficient in the replication is much larger than the coefficient in the 2014 study. However, the fact that this is the only model specification that largely replicates the positive and statistically significant results in the 2014 report (please see sensitivity analyses above) leads us to believe that the main reason for the difference in results because of differences in methodology and not to differences in the sample construction or variables used in the analysis.

[59] Results in Exhibit A41 reflect the results from the LIML instead of 2SLS because it is thought to be a better approach if instruments are weak. The results for the 2SLS were similar.

Exhibit A40

Replication Results

Subject	Grade	Original results from the 2014 report			Results using 2022 report cohorts		
		Coefficient	Standard error	p-value	Coefficient	Standard error	p-value
Math	3	0.137	0.075	0.000	0.545	0.274	0.047
Math	4	0.160	0.076	0.035	1.604	0.288	0.000
Math	5	0.160	0.081	0.047	1.514	0.301	0.000
Reading	3	0.170	0.0071	0.016	0.053	0.270	0.845
Reading	4	0.257	0.094	0.006	0.658	0.253	0.009
Reading	5	0.228	0.103	0.027	0.530	0.273	0.053
Writing	4				0.959	0.288	0.001

We find that these results were not sensitive to the inclusion of parent characteristics, which were excluded in the original study and that the CEM are not sensitive to the inclusion of distance from the ECEAP center (see above). We then looked further into the strength of the IV to see if our reservation about distance satisfying the first criteria for a valid instrument was founded. Exhibit A41 shows the post estimation results from a two-stage least squared regression for math scores over time. This tests the strength of an instrument.

Exhibit A41

Post Estimation Results

Subject	Grade	Robust F-stat	Adjusted R-squared	Partial R-squared	N
Math	3	76.1909	0.01	0.0012	64,448
Math	4	74.4499	0.0135	0.0012	64,231
Math	5	80.9104	0.0143	0.001	63,647
Math	6	63.902	0.0138	0.001	61,784
Math	7	41.1793	0.014	0.0007	60,278
Math	8	37.126	0.0144	0.0006	59,219
Math	10	13.5059	0.0136	0.0008	16,407
Math	11	14.288	0.0128	0.0014	10,069

The literature tends to agree that an F-stat above 10 is a reasonably strong instrument. In all the listed results, the F-stat is above 10, suggesting that the first assumption of the IV holds. However, there is a potential issue with this assumption in the later years, as the F-stat decreases.

Unfortunately, we do not have a way to test for omitted variable bias (that the distance from an ECEAP center affects test scores only through ECEAP attendance). We feel that the hypothetical reasons for the instrument either suffering from omitted variable bias or capturing the results for only a subset of the population made above is compelling.

Conclusion

Ultimately, our partial replication suggests that there are some empirical questions about the validity of the instrument used in the 2014 study. In addition, there are theoretical problems with the methodology which are largely untestable based on the data.

We believe that the updated results are preferable because they do not suffer from bias induced by the IV method. In other words, the estimated program effects from CEM are less likely to be exaggerated and are likely closer to the truth than results from the IV.

However, as mentioned, we cannot claim causal effects even using the CEM methodology. While the method CEM can control for observed characteristics, it is possible that unobserved characteristics in families and unobserved choices by the program itself could drive some of these results. The causal impact of ECEAP may be larger than the observed correlation (the program will have a greater impact than depicted in this report) if the unobserved characteristics are positively related to ECEAP enrollment and negatively related to later outcomes (e.g., it is possible that ECEAP children have more complex needs, after controlling for observed characteristics. They would have performed worse than the comparison group had they not participated in ECEAP). Conversely, the causal impact of the program may be smaller than the observed correlation (the program will have a smaller impact than depicted in this report) if the unobserved characteristics are positively related to ECEAP enrollment and positively related to later outcomes (e.g., It is possible that parents who choose to enroll children in ECEAP are more concerned about academic performance, which would not be captured in the observable data. The children may have performed better than the comparison group even if they hadn't participated in ECEAP). The simplicity of the method allows us to speak more coherently about observed differences between ECEAP participants and non-participants and allow readers to more easily theorize the reasons for the observed differences. We believe that this transparency gives policy makes a stronger foundation upon which to discuss the impact of the program.

For further information, contact:

Chasya Hoagland at 360.664.9084, Chasya.Hoagland@wsipp.wa.gov

Document No. 22-01-2202

Suggested citation: Hoagland, C., Ingraham, B., & Fumia, D. (2022). *Evaluation of the Early Childhood Education and Assistance Program: Short- and long-term outcomes for children* (Document Number 22-01-2202). Olympia: Washington State Institute for Public Policy.