Oregon Childhood Program Participation and Early Educational Outcomes: Part 2 – Risk and Protective Factors

Methods Summary

Table of Contents

Introduction

The inaugural analyses for the Oregon Child Integrated Dataset (OCID) examines relationships between participation in select public early-childhood programs and subsequent early-educational outcomes through 3 interconnected components. *Part 2 – Risk and Protective Factors* examines which child and family characteristics are potential risk and protective factors associated with early-educational outcomes. This summary provides technical documentation of the analysis methods for *Part 2*. The findings from this analysis, as well as more information on the OCID project, are available on the OCID website (www.ocid-cebp.org).

The Center for Health System Effectiveness conducted this analysis in partnership with the Center for Evidence-based Policy, both at Oregon Health & Science University, Portland, OR.

Methods

Study Population

This analysis followed a single academic cohort through the course of their elementary school years, ensuring that any variation in results between outcomes was not driven by changes in the underlying study population composition. As illustrated in Figure 1, the study population attended kindergarten during the 2013-14 school year, which was the first year the state administered the Oregon Kindergarten Assessment. These students attended 5th grade during the 2018-19 school year, corresponding to the latest education data currently in OCID.

The analysis only included children that were born in Oregon, because the OCID population is anchored to Oregon birth records – allowing for examination of associations between circumstances at birth and early-educational outcomes. Additionally, children in the study population attended Oregon public schools, because OCID does not include education information for students who are homeschooled or attend private schools.

Figure 1. Study Population and Points of Outcome Measurement

SCHOOL YEAR	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
GRADE	KG	1st	2nd	3rd	4th	5th

Notes. Green highlighted school year/grade combinations indicate points when outcomes were measured.

Child and Family Characteristics

This analysis assessed a variety of child and family characteristics as potential risk and protective factors associated with early educational outcomes. Characteristics included sex, race/ethnicity (see subsection below), weight for gestational age, prenatal tobacco use, location of birth, mother's age and highest education at the time of birth, English proficiency during kindergarten, enrollment in Medicaid/Children's Health Insurance Program (CHIP), family receipt of income-based supports through Supplemental Nutrition Assistance Program (SNAP) or Temporary Assistance for Needy Families (TANF), and child welfare involvement based on substantiated maltreatment or a foster care placement.

Each measurement of association was made in the context of a comparison group. For example, to understand the association between an educational outcome and being male, comparison was made between the same educational outcome among female students. When serving as covariates in the analysis of other characteristics, additional details about characteristics were included, for example, raw birthweight (in grams) and estimated gestational age (in weeks) were used instead of the rolled-up "small," "average," or "large" weight for gestational age categories. Models also controlled for whether students participated in an Individualized Education Program/Individual Family Service Plan (IEP/IFSP).

For a complete list of child and family characteristics assessed, along with additional information about how they were used in the analysis, please see Appendix A.

Use of Race and Ethnicity Data Across OCID Data Sources

Assigning a race and ethnicity to individuals when there are multiple sources of information across integrated datasets such as OCID is an evolving area of development. Inconsistencies across and within agencies pre-existed OCID in terms of data sources for race and ethnicity information, how agencies collected the data, which categories were available on enrollment and administrative forms, and who responded on behalf of the children for whom there is information. Therefore, the OCID team has been working with state partners to develop a methodology for determining race and ethnicity attributions when there are multiple options within and across points in time. The initial methodology is incorporated in all four components of this OCID Analysis. More information on the calculations can be found in the technical dictionary on the OCID website.

OCID uses race and ethnicity information in dashboards and analyses to describe disparities in outcomes among groups of individuals living in Oregon. Any association between race and ethnicity and outcomes does not imply that the social constructs of race and ethnicity caused that outcome or that there is a biological basis for differences between groups. Personal characteristics such as race and ethnicity often serve as proxies for experiences that are associated with different outcomes for subgroups, such as systemic racism, stress, poverty, or housing instability.

Educational Outcomes

This analysis included 5 early educational outcomes:
- Kindergarten Approaches to Learning
- 2nd Grade Regular Attendance
- 3rd Grade Reading Comprehension
- 5th Grade Student Homelessness
- 5th Grade School Suspensions

Ranges of possible values varied by outcome: Kindergarten Approaches to Learning is measured along a continuum of 1 to 5, with higher scores representing teacher ratings of more advanced

interpersonal skills, self-regulation, and approaches to learning.[1] The remaining 4 outcomes (2nd Grade Regular Attendance, 3rd Grade Reading Comprehension, 5th Grade Homelessness, and 5th Grade Suspensions) are binary, meaning that a student either experienced the outcome or did not.

Calculations followed business rules in use by Oregon's Department of Education at the time of the analysis. See Appendix B for more detailed information about outcome definitions.

Modelling Approach

Associations between child/family characteristics and educational outcomes were measured in terms of average treatments effects (ATEs), after controlling for other observed characteristics. An ATE represents the change in the outcome associated with having the characteristic versus not having the characteristic. Overlap-weighting ensured that comparisons were made between children who could reasonably fall into either category (i.e., having the characteristic or not).

ATEs were modelled using Causal Forests, an extension of Random Forests. This is a statistical approach particularly well-suited for high-dimensional data with complicated relationships between variables. This approach accommodated missing covariate values, which occurred infrequently in our data. Analysis methods were "doubly-robust" because we first estimated the likelihood of having the characteristic ("propensity score"), and then estimated the association between the characteristic and the outcome. Studies have shown this approach yields more accurate (i.e., less biased) results than other common modelling approaches.[2] Models were clustered at the school level to control for correlation in student outcomes between children attending the same school.

As a sensitivity analysis to confirm observed results were not driven by the models' handling of missing covariates, the team re-ran the same models while excluding any students for whom not all variables were observed.

The analysis utilized R, version 3.6.2,[3] and the R package "grf," version 1.2.0.13.[4]

Study Limitations

This analysis is subject to a number of limitations, including:

- Results indicate associations between child/family characteristics and educational outcomes, but do not identify causal relationships or pathways. Furthermore, our findings do not indicate how to intervene, or whether existing intervention is effective.

---

[1] Gill C, Patterson J, Oregon Department of Education, Office of Teaching Learning and Assessment. Oregon kindergarten assessment specifications 2019-2020. 2018; https://www.oregon.gov/ode/educator-resources/assessment/Documents/asmtkatestspecs.pdf. Accessed August 19, 2020.

[2] McConnell KJ, Lindner S. Estimating treatment effects with machine learning. *Health Serv Res.* 2019;54(6):1273-1282. doi: 10.1111/1475-6773.13212.

[3] R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* 2020; www.R-project.org/. Accessed December 2, 2020.

[4] Tibshirani J, Athey S, Friedberg R, et al. Generalized random forest. 2018; https://github.com/grf-labs/grf. Accessed December 2, 2020.

- Results rely on the accuracy and availability of administrative data, which were not collected for the purposes of conducting this analysis.
- Unmeasured factors (e.g., environmental, societal) that were not accounted for in this analysis might help explain an outcome of interest (i.e., "residual confounding"). Our modeling approach aimed to mitigate the bias of these unmeasured factors.
- Although this analysis included a variety of key early-educational outcomes, it does not account for the full complexity of student achievement or well-being.
- This analysis was limited to children who were born to Oregon residents, because the OCID population is anchored to Oregon birth records. This dataset structure allowed us to examine associations between circumstances at birth and early-life educational outcomes but prevented inclusion of Oregon public school students that were born out-of-state. The OCID population includes approximately three-quarters of all enrolled Oregon public school students, although this completeness varies by demographic characteristics (e.g., gender, race/ethnicity, and age-grade-level).[5] Additionally, OCID does not include education information for students who are homeschooled or attend private schools.
- When assessing family characteristics, children were linked to the parents identified on their birth record; these parents might not be the residential or legal guardians of the child.
- Some of characteristics included in this analysis were time-varying (e.g., family receipt of SNAP, TANF, and Medicaid/CHIP benefits are broken out by months of enrollment per year, for each of the first 5 years of the child's life) whereas others were measured at a point in time (e.g., mother's education at birth). Point-in-time measurements offered a snapshot of a child or family characteristic at a particular phase in life, but did not allow us to account for changes in the characteristic over the child's life course. For more information about each of the child and family characteristics, and how they were used as covariates, see Appendix A and the footnote corresponding to Table A1.
- Some outcomes included in this analysis rely on subjective ratings of student abilities (e.g., Kindergarten Approaches to Learning) and behaviors (e.g., 5th Grade Suspensions). Use of school-level clusters did not adjust for correlation among outcomes driven by a particular individual (e.g., teacher or principal).
- This analysis included multiple statistical tests. Through sequential testing of each characteristic-outcome pair, the probability of falsely identifying an association as significant increases.

---

[5] Center for Evidence-based Policy. Explore the OCID population. 2020; https://www.ocid-cebp.org/about-ocid/explore-ocid/. Accessed August 26, 2020.

Acknowledgements

Appendix A. Child and Family Characteristics

Table A1. Sources and Levels of Child and Family Characteristics

CHARACTERISTIC	DATA SOURCE	LEVELS (COMPARISON GROUP BOLDED)
Gender	ODE	– **Female** – Male
Race/ethnicity [a]	Multiple program sources: DHS (AFCARS/NCANDS), EI/ECSE, ODE, OHA, OPK, and home visiting programs	– American Indian/Alaskan Native – Asian – Black/African American – Hispanic/Latinx – Multiracial – Native Hawaiian/Pacific Islander – **White**
Weight for gestational age [b]	Birth record	– Small for gestational age (SGA) – **Average for gestational age (AGA)** – Large for gestational age (LGA)
Prenatal tobacco use	Birth record	– **No prenatal tobacco use** – Prenatal tobacco use
Mother's education at birth	Birth record	– Less than high school – **High school graduate/GED** – Some college – College graduate – Advanced degree
Mother's age at birth	Birth record	– <20 years – **20-29 years** – 30+ years
Geography	Birth record	– Tri-County – **Urban** – Rural
Limited English proficiency [c]	ODE	– **None** – Some (any duration)
SNAP enrollment [d]	ICS	– **None** – Some (any duration)
TANF enrollment [d]	ICS	– **None** – Some (any duration)
Medicaid/CHIP enrollment [d]	OHA	– **None** – Some (any duration)
Substantiated maltreatment [e]	DHS (NCANDS)	– **None** – Some (any duration)
Foster care placement [f]	DHS (AFCARS)	– **None** – Some (any duration)

Notes: [a] Race and ethnicity information from multiple program sources was used, and allowed identification to change over time and prioritizing self-identification of child or caregiver from sources that offered disaggregated categories. [b] Birth weights less than the national 10th percentile (in grams, adjusted for gender and gestational age) and greater than the national 90th percentile (same adjustments) are considered small and large, respectively, based on U.S. 2009-10 live birth files.[18] When used as covariates (to adjust for other exposures), raw birthweight and estimated gestational age are used instead of the rolled-up "weight for gestational age." [c] Determined at kindergarten. [d] When used as covariates (to adjust for other characteristics), family receipt of SNAP, TANF, and Medicaid/CHIP benefits are broken out by months of enrollment per year, for the first 5 years of the child's life. [e] When used as a covariate (to adjust for other exposures), maltreatment is broken out by broad category: neglect, medical neglect, physical, psychological/emotional, and sexual. [f] When used as a covariate (to adjust for other exposures), foster care placements are broken out as an indicator for each of the first 5 years of the child's life. Abbreviations. AFCARS: Adoption and Foster Care Analysis and Reporting System; DHS: Department of Human Services; EI/ECSE: Early Intervention and Early Childhood Special Education; GED: general education development certificate; ICS: Integrated Client Services; NCANDS: National Child Abuse and Neglect Data System; ODE: Oregon Department of Education; OHA: Oregon Health Authority; OPK: Oregon Pre-Kindergarten.

Appendix B. Early Educational Outcomes

Name: **Kindergarten Approaches to Learning**

Definition: Teacher-rated Kindergarten Approaches to Learning values represent an average score across the 15-item Child Behavior Rating Scale (CBRS). Results fall along a development continuum from 1 to 5, with higher scores representing a more advanced demonstration of interpersonal skills, self-regulation, and approaches to learning. This is the only outcome in the study that was measured on a continuous scale.

Name: **2nd Grade Regular Attendance**

Definition: Students who attend greater than 90% of school days during an academic year. Students must be enrolled for at least 75 instructional days in an academic year, consistent with Oregon Department of Education business rules. In the event of school transfers, present and absent days are summed across enrollment segments. This was a binary outcome; children either had regular attendance or did not.

Name: **3rd Grade Reading Comprehension**

Definition: Students who meet or exceed grade-level reading standards. During the 2016-17 school year, this corresponds to achievement of level 3 ("meets") and level 4 ("exceeds") on the Oregon Smarter Balanced Assessment. This was a binary outcome; children either met/exceeded grade-level reading standards or did not.

Name: **5th Grade Homelessness**

Definition: Students who are determined to be McKinney-Vento eligible for any duration of an academic year. This designation includes several living situations: Shared housing ("doubled up"); Shelter; Motel/hotel; and Unsheltered. This was a binary outcome; children either experienced homelessness or did not.

Name: **5th Grade Suspensions**

Definition: Students who are suspended one or more times during the school year. Includes in-school and out-of-school suspensions. This was a binary outcome; children were either suspended or were not.