Who attends public pre-kindergarten? A model-based approach to understanding similarities and differences among Michigan public pre-K students

Jamie Heng-Chieh Wu, Ph.D. & Hope O. Akaeze, Ph.D.

Michigan State University

Suggested citation: Wu, J. H., & Akaeze, H. O. (2025). Who attends public pre-kindergarten? A model-based approach to understanding similarities and differences among Michigan public pre-K students. [White Paper]. Community Evaluation Programs at University Outreach and Engagement, Michigan State University.

Acknowledgments: The study was supported by the Michigan Department of Education through the State Evaluation of the Great Start Readiness Programs grant (award number: 215081-GSRPEV6).

Abstract

Studies of publicly funded pre-kindergarten (pre-K) programs typically rely on descriptive statistics, using basic demographic data to understand program participants. Using descriptive summaries, these studies create limited descriptions based on one or two demographic characteristics at a time, such as race/ethnicity and income. Using rich data on student risk factors collected for Michigan's state-funded pre-K, the Great Start Readiness Program (GSRP), we conducted a model-based analysis of program participants. Our latent class analysis model revealed five distinct classes among participants based on risk factors the Michigan Department of Education uses to determine eligibility for GSRP. Rather than defining participants with simplistic labels based on family income or race/ethnicity, our approach yields a more nuanced portrait of GSRP students and of the needs and challenges GSRP families face. Equipped with richer information about the profile of pre-K families, policymakers and program administrators can target program services to meet the specific needs of groups of children and families.

Keywords: Public pre-kindergarten, preschool, student demographics, latent class analysis (LCA), student profile

Studies of the student populations of publicly funded pre-kindergarten (pre-K) programs generally rely on basic demographic statistics on family income, race/ethnicity, gender, and sometimes English language learner status. Using descriptive summary approaches, these studies either provide descriptive statistics for reports to funders or correlate one or two variables at a time with the outcomes of interest. Students are grouped based on limited and predetermined characteristics, such as Black low-income vs. White middle-income students. Then these broadly defined groups are compared on outcomes of interest.

By contrast, model-based approaches, such as latent class analysis (LCA) or latent profile analysis (LPA), correlate many variables to create nuanced clusters of students based on observed characteristics (Bartholomew et al., 2011; Oberski, 2016). Most of the studies of pre-K students using LCA and LPA rely on data collected in kindergarten, and most construct classes of students in order to examine how class membership correlates with academic and behavioral outcomes (e.g., Christensen et al., 2022; Helsabeck et al., 2021; Rhoad-Drogalis et al., 2020).

Our study is the first to use rich data on the background characteristics of students in a state-funded pre-K program to develop profiles of student risk factors using LCA. This research joins an LCA study conducted on Head Start students using a different set of background characteristics (Rhoades Cooper & Lanza, 2014) in helping the field understand which children from which backgrounds attend publicly funded early childhood centers. LCA enables researchers to create detailed participant profiles that go beyond simple classification by race/ethnicity and family income levels. This model-based approach facilitates examination of demographic and family risk factors among persons, capturing in part the multidimensional nature of students' identities. While descriptive summaries present data at a global level and apply labels—for example, noting that a certain percentage of participants are "black and low income" —the model-based LCA approach taken in our research enables us to more fully understand which families, with which characteristics or combination of risk factors, are using state-funded pre-K programs. It enables us to explore the nuances of families' unique circumstances and unmet needs beyond the over-simplified label of "low-income families." Because Michigan's public pre-K gives enrollment priority to students based on risk factors that can affect their ability to succeed in school in the future, LCA helps to reveal intersections among those risk factors and standard demographics such as race/ethnicity and income. The LCA results can help policymakers better understand who is being served by public pre-K programming. Because participation in public pre-K is voluntary, the analysis also shows the extent to which families with varying risk levels and demographic backgrounds are interested in enrolling their children in public pre-K. Our findings can help policymakers and early childhood administrators tailor services to meet the varying needs of different groups of children.

Who Is Eligible to Enroll in Michigan's Public Pre-K Program

This study is possible because of the unique nature of the enrollment policy implemented by the Michigan Department of Education (MDE) in its public pre-K program, the Great Start Readiness Program (GSRP). Like 34 of the 44 states nationwide that provide public pre-K programs (Friedman-Krauss et al., 2022), Michigan offers priority in enrollment to low-income families. Families are sorted by percentage of FPL into quintiles: 0–50% of FPL, 51–100%, 101–150%, 151–200%, and 201–250%. Having one of three additional eligibility factors automatically places children in the lowest quintile, regardless of actual income: if the child has a qualifying IEP, is experiencing homelessness, or is in the foster care system. Six additional eligibility factors serve as "tie breakers" for families who have the same percentage of FPL:

disability, abuse or neglect, home language other than English, severe challenging behavior, environmental risk, and low parental education (Michigan Department of Education, 2020). To determine which children to admit to the program, intermediate school districts (ISDs, the administrators of GSRP grants) start with the lowest-income families. If two families have the same percentage of FPL, the one with more eligibility factors is admitted first. After all applying children in the lowest quintile are enrolled, children in the next-lowest quintile are considered.

Children whose family income is above 251% of FPL may be admitted after all children who qualify on the basis of income have been admitted. Over-income families pay a sliding-scale fee determined by the ISDs. During the 2020–2021 school year, in anticipation of lower enrollment due to the COVID-19 pandemic, MDE increased the income limit to 400% of FPL and removed the cap on the percentage of over-income families an ISD could enroll. In previous years, such families could make up only 10% or less of the ISD's total GSRP enrollment. In the 2021–2022 school year, the cap was reinstated but increased, so that up to 15% of enrollment could consist of over-income children.

Our review of NIEER's *State of Preschool* annual report (Friedman-Krauss et al., 2022) suggests that GSRP is unique among state pre-K programs in the number of risk factors it considers to supplement the primary eligibility factor, family income. In addition to standard demographics on gender, race/ethnicity, and income, Michigan collects data on nine widely varying family and child factors, all of which are known to affect children's educational outcomes. Our access to this rich trove of data, as the contracted evaluators of GSRP, enabled our study of student characteristics beyond income and race/ethnicity.

Approaches to Profiling Pre-K Students

Descriptive Summary Approach

Descriptives on student characteristics, taken individually, give an overview of the populations served by public pre-K programs. In Michigan's GSRP, about 95% of participants every year come from low-income families— by design, because these families receive priority in enrollment. Almost half of participants are subject to environmental risk, defined by MDE as loss of a parent due to death, divorce, incarceration, military service, or absence; teen parent; homelessness; residence in a high-risk neighborhood; or pre- or postnatal exposure to toxic substances (Michigan Department of Education, 2020). This information on specific characteristics gives policymakers a high-level view of the extent to which the programs reach the targeted populations and help address the opportunity gaps caused by poverty. Hundreds of studies, including evaluations of state- and city-funded pre-Ks, have used similar descriptive summary approaches (e.g., Durkin et al., 2022; Gormley et al., 2018; Gray-Lobe et al., 2021). However, the descriptive summary approach is limited by its ability to identify subgroups based on only one or two characteristics (Collins & Lanza, 2009; Lanza & Rhoades, 2013). It assumes a one-dimensional relationship among factors, so that, for example, coming from a low-income family is associated with a high probability of environmental risk and of all other risk factors. Researchers may compensate for this limitation by pre-setting a limited combination of variations for analysis, such as combining income levels with English language status. Even so, this approach limits understanding of participants' identities to a few pre-determined factors and forestalls exploration of more nuanced views of participants' identities.

Model-Based Approach

The model-based approach embodied in LCA offers a fuller understanding of which families are served by publicly funded pre-K programs, revealing the existence of subgroups whose characteristics vary in their interactions (Collins & Lanza, 2009; Lanza & Rhoades, 2013). For example, families who speak a language other than English at home have a broad range of income levels, so that ELL status and low income are not always correlated, even though both ELL status and low income are educational risk factors. LCA permits researchers to categorize children along multiple dimensions to arrive at a nuanced portrait of interacting identities and risk factors. It therefore can enrich understanding of the lived experiences of program families, including the challenges they face if their children have one or more risk factors. Policymakers and program administrators can use this understanding to provide more targeted and effective services based on the needs of identified groups of children.

Several studies in the education literature have demonstrated the utility of LCA and its appropriateness for creating typologies of young children (e.g., Helsabeck et al., 2021, Nylund et al., 2007). Justice and colleagues (2017), for example, found four distinct classes of kindergarten readiness among low-income children in Appalachia who had attended public pre-K programs. Other studies have used LCA to correlate profiles of child participation (or lack of participation) in center-based care with kindergarten readiness (e.g., Helsabeck et al., 2021; Rhoad-Drogalis et al., 2020). Christensen and colleagues (2022) constructed profiles of entering kindergarteners based not only on child characteristics but also on risk factors related to family and social background characteristics. The seminal work by Kim and Fram (2009) on parents' priorities in their child care choices also used LCA.

We have found only one other study that uses LCA to identify classes of children based on risk factors they bring into the early childhood setting. A study of Head Start students (Rhoades Cooper & Lanza, 2014) identified classes of students based on the home and maternal characteristics targeted in the Head Start impact study (Puma et al., 2010); it then correlated those classes with differential effects of Head Start participation at various points through first grade.

Methods

Data

Data for this analysis comes from the MDE GSRP database, covering five cohorts of GSRP participants, from 2017–2018 to 2021–2022. Use of multi-cohort data enables a more robust participant profile than would be possible with only one year of data and reveals any changes over time. The total number of participants for the five years was 178,232 (see Table 1). In the first three years, the number of participants hovered near 38,000. Enrollment fell during the first full COVID year, 2020–2021, to about 28,400. In 2021–2022, enrollment nearly bounced back to pre-pandemic levels, at almost 36,500.

We also used data on the racial/ethnic composition of the Michigan population as a whole from the U.S. Census Bureau American Community Survey 2020, whose five-year average covers roughly the study period. For comparative purposes, we report on the racial/ethnic makeup of the population of Michigan children whose family income, at 250% or less of FPL, makes them eligible for GSRP. These data were derived from MDE school records covering kindergarteners in 2017–2021, whose demographics are presumed to be nearly identical to those of children one year younger.

Analytical Approach

The variables we analyzed are seven eligibility criteria for GSRP enrollment. We analyzed data on family income as a percentage of FPL and six binary variables: whether the child has a disability or developmental delay, experiences abuse or neglect, does not speak English at home, has severe challenging behavior, or is exposed to environmental risk, and whether the child's parents have low educational attainment. LCA is appropriate for person-centered clustering where observed variables are categorical and latent variables are believed to be categorical (Bartholomew et al., 2011). All observed variables in our data are categorical, except for the FPL, which is ordinal. Although family income exists as a continuum, we followed MDE in categorizing percentages of FPL into quintiles: 0–50%, 51–100%, and so on up to 301% or more. We did not analyze the other three eligibility criteria in Michigan's pre-K enrollment policy—whether the child is in foster care, is homeless, or has a qualifying IEP—because the policy equates those criteria with low family income.

All analyses were carried out using Mplus version 8.4 (Muthén & Muthén, 1998-2017). One advantage of LCA is that researchers can assess model fit, using statistics that result from implementing LCA in most popular statistical software packages, to enhance objectivity and accuracy in model selection (Collins & Lanza, 2009; Lanza & Rhoades, 2013). Model choice was based on comparing the fit statistics for model solutions with 2 to 6 classes, using standard fit statistics: AIC, BIC, SBIC, chi-square, and entropy statistics. Another consideration was model interpretability: the extent to which the resulting classes were understandable and could be differentiated from one another.

Results

Descriptive Summary Results

The annual evaluation reports we produce for MDE exemplify the descriptive summary approach to studying program participants. The reports include breakdowns of the student population by income (Table 1), risk/eligibility factors (Table 2), and race/ethnicity (Table 3). The breakdowns are fairly consistent for the first three years. Changes emerge in 2020–2021, when MDE removed the cap on the percentage of over-income families who could enroll. In 2021–2022, MDE reinstated the cap on over-income families but increased it from the pre-pandemic level of 10% to 15%. Those decisions had the expected effect on enrollment in 2020–2021 and 2021–2022, skewing the proportion somewhat toward families who would not otherwise have been eligible, at more than 251% of FPL (Table 1).

Table 2 shows the distribution of risk factors that, together with income, determine children's eligibility for GSRP. After income, the second largest category, year after year, is environmental risk, which encompasses parental loss (including through divorce or military service, among others, as well as death), teen parent, unstable housing, high-risk neighborhood, and exposure to toxic substances.

Table 3 shows that the racial/ethnic makeup of the GSRP population has been relatively consistent during the five years of the study, with the exception of the 2020–2021 school year, when MDE removed the income cap in response to the COVID-19 pandemic. In that year, the percentage of Black children was somewhat lower than usual, while the percentage of White children was somewhat higher. Proportions of other racial/ethnic categories remained about the same. In 2021–2022, by contrast, the proportion of Black children was slightly higher and the proportion of White children slightly lower, with other groups remaining about the same.

These demographic and risk factor data help program administrators and policymakers to determine the extent to which GSRP is reaching its intended populations. A descriptive summary approach to profiling the pre-k children thus has its function as a policy tool.

Model-Based Results

The model-based LCA approach, which we applied to the GSRP eligibility factor data after providing the basic descriptive summary outlined above, resulted in a much richer and more nuanced picture of GSRP children and families. Application of LCA produced and examined models ranging in size from 2 classes to 6. The entropy statistics were strongest for the 4-class solution, but the other fit statistics preferred the 5- and 6-class solutions. The 6-class solution presented better fit statistics than the 5-class solution, but the 5-class model had better class interpretability. We therefore used the 5-class solution.

LCA Findings

Table 4 presents the distributions of the factors that determined enrollment priority, namely, the income level and eligibility factors, in each of the 5 classes. **Class 5,** *extreme low income, high risk,* consists of the most vulnerable children. Their families have extremely low FPL percentages, and the class presents high or moderately high percentages in almost all eligibility factors except home language other than English. **Class 4,** *extreme low income, selective risk,* encompasses typical GSRP participants. Their families have extremely low incomes, but the children are less likely to be exposed to abuse and neglect, to demonstrate severe challenging behaviors, or to have a disability. This group has the highest rate of home language other than English and the second-highest rates of parents with low education and of environmental risk. This class consistently represents more than 40% of the total GSRP population, year after year. **Class 3** is *low income, average risk*. **Class 2** is *low income, lower* *risk* compared to Class 3. **Class 1** is *higher income, selective risk;* these children's families are over-income for GSRP, but the children may have other risk factors, particularly disability, abuse and neglect, and/or severe challenging behavior. The use of the term "higher income" here is deliberate in the sense that **Class 1** has higher income relative to the rest of the GSRP population, not that they are considered as higher income in comparison to the general Michigan population.

Compared to the differences among the other classes, Class 4 and Class 5 are quite similar in terms of family income. However, they are quite different in terms of other risk factors. For example, Class 4 has the lowest percentage among all the classes of children with severe challenging behavior and of children who have experienced abuse and neglect, while Class 5 has the highest percentage of children in those two categories. Conversely, Class 4 has the highest percentage of children whose home language is not English while Class 5 has the lowest. Equally striking are the differences between Class 1, over-income children, and Class 4, the largest class of typical GSRP students. In almost all categories, Class 1 is a mirror image of Class 4. Class 1 has the highest or second-highest proportion of students with disabilities, severe challenging behavior, and abuse and neglect, while Class 4 has the lowest proportion of students in those categories.

Because GSRP enrollment priority is largely based on income, and risk factors are considered only when families have the same income levels, the order of the five classes largely reflects the weighing of the enrollment priorities, with Class 5 being the top priority group and Class 1 likely to be absent unless there are seats available after programs accommodate students from all other groups. Figure 1 displays the prevalence of these classes in all five cohorts of students. Class 4 is largest, and Class 1 has the smallest number of children. Over the past five years, approximately half of GSRP participants have fallen into Classes 4 and 5. Class 1, children from relatively higher-income families who may be developmentally at risk because of disability, behavioral challenges, or abuse and neglect, made up about 6% of the GSRP population in the past five years.

Figure 2 depicts the information from Table 4 in a graph. The vertical axis represents the conditional probability of each risk factor—that is, the probability of that factor being present in a child belonging to the given class. The FPL description was included in the legend but not in the line graph because it is different in scale from the other risk factors. We added the three eligibility factors that were not included in the LCA as the last three points on the horizontal axis of the graph in order to enrich the profiles of the classes. These three extra eligibility factors tend to support the class divisions: Class 5 children are at highest risk for all three factors, and children in Classes 2, 3, and 4 fall in the middle, although Class 4 children have a slightly higher risk of homelessness. As expected, Class 1 children have low percentages on the foster care and homelessness factors but have a relatively high proportion of qualifying IEPs.

Because the enrollment rules changed starting in 2020–2021 to enable ISDs to enroll a higher proportion of children whose family income was greater than 250% of FPL, we also examined the class distributions for each of the five years separately. Figure 3 shows that enrollment overall was lower in 2020–2021 than in the previous years, in keeping with national trends documented by Friedman-Krauss et al. (2022). It also shows that the distribution of classes was stable in the first three school years but changed in 2020–2021 and 2021–2022. As expected, the proportion of children in Class 1, *higher income, selective risk,* was higher in 2020–2021—nearly double the average in school years 2017–2018 through 2019–2020. In 2021–2022, the proportion of children in Class 1 was higher than in the first three years but lower than in 2020–2021. The variations in Classes 2, 3, and 5 between pre-pandemic and pandemic years were smaller. A meaningful difference emerged in Class 4, which had a lower proportion of children in 2020–2021 and 2021–2022 than in previous years—although, as Figure 3 suggests, this trend may have started in the 2019–2020 school year. Class 4 represents typical GSRP participants, who qualify on the basis of their extremely low family income and have widely varying levels of other risks, as shown in Table 4.

Findings from Overlay of Racial/Ethnic Data

To deepen understanding of the experiences and identities of GSRP children and their families, we explored the racial/ethnic composition of the five LCA classes. For comparison, Figure 4 shows the racial/ethnic breakdowns of the population of Michigan and of the income-eligible population, that is, children of families whose income is 250% or less of FPL. Figure 5 shows the racial/ethnic composition of the five LCA classes. Both figures collapse the multiracial, Asian, Pacific Islander, and American Indian categories into the single category "Other" because the proportions of these groups in the GSRP population are too small to permit analysis.

Figure 4 shows that Michigan as a whole is predominantly White. In the GSRP-eligible population, White students still make up the majority, but Black and Hispanic students are more strongly represented.

In Figure 5, the racial/ethnic makeup of the income-eligible population from Figure 4 has been superimposed as transparent bars behind the solid bars representing racial/ethnic makeup of the GSRP participants in each class. The striking finding is the predominance of Black children in Class 4, the largest class comprising "typical" GSRP children. Black children constitute 40% of Class 4, as compared to 29% of the income-eligible population. White children, by contrast, are 41% of Class 4, though they make up 52% of the income-eligible population. In all other classes, the disproportion is reversed: Black children make up a smaller proportion of Classes 1–3 and Class 5 than their proportion of the income-eligible population, while White children make up a larger proportion of those classes than their proportion of the income-eligible population. Less striking discrepancies emerge among Hispanic children, whose proportions in Classes 3 and 4 match those of their proportion in the income-eligible population but are a little lower in Class 2 and noticeably lower in Classes 1 and 5. The proportion of GSRP participants in the "other" group corresponds almost perfectly to the state's income-eligible population in that group.

Discussion

In our model, sorting by the factors used to determine GSRP eligibility reveals an uneven distribution of risk factors and racial/ethnic backgrounds among the five classes. These differences have implications for policy and practice.

LCA Findings

In the LCA findings, a strong set of contrasts emerges between Class 1, over-income children, and Class 4, the largest class comprising typical GSRP participants. In terms of the primary eligibility factor, family income, the two classes are, by definition, nearly at opposite ends of the spectrum. On almost all other eligibility factors, where Class 1 is high, Class 4 is low and vice versa. Class 4 has the lowest rates of disability, severe challenging behavior, and abuse or neglect, while Class 1 has the highest or second-highest rates in these categories. Conversely, Class 1 has the lowest rates of low parental education and environmental risk, while Class 4 has the second-highest rates. Class 1 also has the second-lowest rate of home language other than English (after Class 5, the lowest-income class) and Class 4 has the highest.

Even in their proportion of enrolled students during the worst year of the pandemic, Class 4 fell while Class 1 rose. The increase in the proportion of Class 1 students is likely a direct result of the policy change that enabled more over-income students to enroll. That policy change was enacted in anticipation of a drop in enrollment due to COVID. What was not anticipated was that, while enrollment fell across the four classes of income-eligible students, the largest proportional drop would be in the largest class, Class 4. Meanwhile, Class 5 children, who were at highest risk in almost all aspects other than home language other than English, and Class 2 or Class 3 children, who had slightly higher income and risk in selective areas, enrolled at similar rates as in previous years.

Another way to explore differences in the classes is to identify discrepancies between the primary eligibility factor, income, and other risk factors. In rates of disability, severe challenging behavior, and abuse and neglect, Class 4 has the lowest rate, and Class 1 and Class 5 have the highest or second-highest rates. Policymakers might assume that Class 5, the lowest income group, would have the highest rate of most risk factors, followed by Class 4, whose income level is only slightly higher than that of Class 5. In fact, Class 1, the most relatively affluent class, has high rates of these risk factors. Only two risk factors, low parental education and environmental risk, track with income, with Class 5 having the highest risk rate and Class 1 the lowest.

Furthermore, home language other than English, one of the determining factors for enrollment priority, seems to have no relationship with income. Figure 2 shows that this risk factor has the smallest level of variation from one class to another, with the incidence clustering in a small range across all five classes. As Table 4 shows, the distribution of home language other than English among the classes shows a different pattern from that of other risk factors. Class 5, which has the highest proportion of other risk factors, has the lowest rate of home language other than English. Class 4, which is just a little different from Class 5 in terms of family income, has the highest rate of home language other than English.

Findings from Overlay of Racial/Ethnic Data

Taken together, GSRP participants roughly represent the racial/ethnic composition of the income-eligible population of Michigan as a whole. Comparison of Table 3 with Figure 4 shows that Black people are 29% of income-eligible individuals and 27% of GSRP enrollees; White children are 52% of the income-eligible population and 54% of enrollees; other groups are also roughly proportional. This finding from a descriptive summary approach contrasts with the finding of a recent equity-focused study of Head Start, in which White and Asian children were underrepresented in comparison to their proportion of the income-eligible population (Friedman-Krauss, Barnett, & Duer, 2022). If anything, White children are slightly overrepresented in GSRP.

When we look at the five classes using a model-based approach, however, differences do emerge in the participation rates of the racial/ethnic groups. The most salient finding is that the racial/ethnic makeup of Class 4—the largest of the five classes, at 43.5% of total enrollment—is substantially different from the makeup of all other classes. As Figure 5 shows, Black children are substantially overrepresented in Class 4, while White children are substantially underrepresented, and the proportions of children identified as Hispanic and "other" are about the same as in the income-eligible population. In all other classes—including Class 5, the lowest-income, highest-need group—Black children are underrepresented and White children are overrepresented.

Policy Implications

Using data on risk factors, overlaid with data on racial/ethnic background, has given us a richer picture of GSRP participants than would have been possible with a straight descriptive summary approach. The broadest finding is that participant profiles have been generally consistent during the past five years. Even in the midst of the pandemic, when enrollment dropped significantly, the proportion of types of students with varying risk factors remained similar. Specifically, Class 4, comprising children whose families have extremely low income, shows relatively low levels of most other risk factors over the last five years. Class 4, by far the largest of the five classes at about 43% of participants, is also the only class in which Black and White children make up nearly equal proportions of the population; White children are underrepresented compared to their proportion of the eligible population, and Black children are overrepresented.

These findings suggest that typical GSRP participants, the ones in Class 4, need educational services designed to boost the academic and developmental readiness of low-income children, but most do not need special services designed for children with other risk factors, such as disability, severe challenging behavior, or abuse and neglect. In fact, policymakers and program administrators should concentrate services for children with disabilities among the children with the highest level of family income, almost 23% of whom have one or more disabilities, and those in the lowest-income class, 22% of whom have a disability. The higher-income class also has high rates of severe challenging behavior and abuse and neglect. The extremely low-income class, which accounts for about 6% of all participants, needs intensive services in all areas except English language learning. Policymakers and program administrators should target high-impact support for children with disabilities, behavior issues, parents with low educational attainment, abuse and neglect, and environmental risk to children in this lowest-income class. Meanwhile, these decision-makers should emphasize typical (Class 4) participants when designing interventions to support English language learners, since 13% of this class comes from a home where the first language is not English.

How deeper understandings of the participants in public pre-K programs might influence broad policy discussions, such as the ongoing debate over making state-funded pre-K available to all children regardless of income, remains to be seen. Our data show that, when more over-income children were allowed to enroll to fill open seats, they did show up—even in the worst year of the pandemic. This surge demonstrates the interest of slightly higher-income families in enrolling their children in GSRP.

To better illuminate policy debates about participation in public pre-K, at minimum, the field needs LCAs of other states' pre-K programs. This observation leads to our main policy recommendation: States ought to collect as much data as is feasible on their pre-K program participants and families. Our LCA on Michigan participants and the LCA on Head Start participants by Rhoades Cooper and Lanza (2014) illustrate the value of collecting data including not only basic demographics but also child and family characteristics known to affect children's school success. Ideally, the federal and state governments would agree to collect the same data. Then meaningful comparisons could be made among publicly funded early education programs nationwide.

Limitations

Our LCA focuses on one state's publicly funded pre-K program. The findings are not likely to be generalizable to other public pre-Ks targeting low-income families. In general, LCAs of different populations can be expected to produce different class distributions. There is no reason to expect that LCA results would be the same in Idaho or New York as in Michigan. In particular, Michigan's unique policy of prioritizing children with risk factors other than low income served as the basis for our LCA; analyses conducted in other states may have less or no data on risk factors. Furthermore, we overlaid data on race/ethnicity onto the classes produced by the LCA of the risk factors defined by GSRP policy. Other studies that incorporate race/ethnicity into the LCA could yield different results.

Future Directions

This LCA study is part of a longitudinal research project into the effectiveness of GSRP in improving academic and behavioral outcomes for low-income children. As school outcome data become available for all cohorts of students, we are eager to examine the correlations between class membership on the one hand and academic and social-emotional outcomes on the other. A substantial body of work has examined the question of which students benefit most from public pre-K—but the findings are mixed. The ability to create nuanced, model-based profiles of student classes may help to explain the differential effects of public pre-K participation on later outcomes.

References

Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). *Latent variable models and factor analysis: A unified approach.* John Wiley & Sons.

Christensen, D., Taylor, C., Hancock, K., & Zubrick, S. (2022). School readiness is more than the child: A latent class analysis of child, family, school, and community aspects of school readiness. *Australian Journal of Social Issues, 57*, 125-143. https://doi.org/10.1002/ajs4.138

Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences.* John Wiley & Sons.

Durkin, K., Lipsey, M., Farran, D., & Wiesen, S. (2022). Effects of a statewide pre-kindergarten program on children's achievement and behavior through sixth grade. *Developmental Psychology*, Advance online publication. http://doi.org/10.1037/dev0001301

Friedman-Krauss, A., Barnett, S. W., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., Gardiner, B. A., & Jost, T. (2022). *State of preschool 2021.* National Institute for Early Education Research, Rutgers Graduate School of Education. https://nieer.org/wp-content/uploads/2022/04/YB2021_Full_Report.pdf

Friedman-Krauss, A., Barnett, W., & Duer, J. (2022). *The state(s) of Head Start and Early Head Start: Looking at equity.* National Institute for Early Education Research.

Gormley, W., Phillips, D., & Anderson, S. (2018). The effects of Tulsa's pre-K program on middle school student performance. *Journal of Policy Analysis and Management, 37*, 63-87. https://doi.org/10.1002/pam.22023

Gray-Lobe, P., Pathak, P., & Walters, C. (2021, May). *The long-term effects of universal preschool in Boston.* School Effectiveness & Inequality Initiative, MIT Department of

Economics. https://seii.mit.edu/wp-content/uploads/2021/05/SEII-Discussion-Paper-2021.4-Gray-Lobe-Pathak-and-Walters.pdf

Helsabeck, N. P., Logan, J. A., Justice, L. M., Purtell, K. M., & Lin, T.-J. (2021). Pathways to kindergarten: A latent class analysis of children's time in early education and care. *Early Education and Development, 32*(5), 764-788.

Justice, L., Jiang, H., Khan, K., & Dynia, J. (2017). Kindergarten readiness profiles of rural, Appalachian children. *Journal of Applied Developmental Psychology, 50*. http://doi.org/10.1016/j.appdev.2017.02.004

Kim, J., & Fram, M. (2009). Profiles of choice: Parents' patterns of priority in child care decision-making. *Early Childhood Research Quarterly, 24*, 77-91. https://doi.org/10.1016/j.ecresq.2008.10.001

Lanza, S. T., & Rhoades, B. L. (2013). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science, 14*(2), 157-168.

Michigan Department of Education. (2020). *GSRP implementation manual. Section: Eligibility.* https://www.michigan.gov/-/media/Project/Websites/mde/gsrp/implementation/gsrp_eligibility_section.pdf?rev=a3fda886d6464d02916ae0a2a855a674

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide.* 8th ed. Muthén & Muthén.

Nylund, K., Bellmore, A., Nishina, A., & Graham, S. (2007). Subtypes, severity, and structural stability of peer victimization: What does latent class analysis say. *Child Development*, 1706-1722.

Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Robertson, *Modern statistical methods for HCI.* Springer. https://doi.org/10.1007/978-3-319-26633-6_12

Puma, M., Bell, S., Cook, R., & Heid, C. (2010). *Head Start impact study final report.* U.S. Department of Health & Human Services, Administration for Children and Families. https://files.eric.ed.gov/fulltext/ED507845.pdf

Rhoad-Drogalis, A., Justice, L., Tzu-Jung, L., Purtell, K., & Logan, J. (2020). Profiles of preschool attendance and children's kindergarten readiness. *Early Education and Development*. https://doi.org/10.1080/10409289.2020.1799618

Rhoades Cooper, B., & Lanza, S. (2014). Who benefits most from Head Start? Using latent class moderation to examine differential treatment effects. *Child Development, 85*(6), 2317-2338. https://doi.org/10.1111/cdev.12278

Tables and Figures

Table 1

Federal Poverty Levels of GSRP Participants, 2017–2022

Year	2017–2018	2018–2019	2019–2020	2020–2021	2021–2022	Totals
Number of students	38,088	38,075	37,232	28,422	36,415	**178,232**
0–50% FPL	30%	30%	27%	25%	27%	**28%**
51%–100% FPL	24%	24%	23%	22%	20%	**23%**
101%–150% FPL	20%	20%	20%	19%	20%	**20%**
151%–200% FPL	13%	13%	14%	14%	14%	**14%**
201%–250% FPL	9%	10%	10%	10%	11%	**10%**
251%–300% FPL	2%	2%	3%	4%	4%	**3%**
301% and above FPL	3%	2%	3%	7%	5%	**4%**

Table 2

Eligibility Factors of GSRP Participants, 2017–2022

Year	2017–2018	2018–2019	2019–2020	2020–2021	2021–2022	Totals
Number of students	38,088	38,075	37,232	28,422	36,415	**178,232**
Low-income (250% FPL and below)	96%	96%	95%	89%	92%	**94%**
Environmental risk	55%	52%	47%	46%	47%	**50%**
Parent low educational attainment	18%	17%	18%	15%	14%	**16%**
Disability	13%	11%	11%	12%	12%	**12%**
Home language not English	9%	9%	9%	10%	10%	**9%**
Abuse or neglect	9%	9%	9%	9%	8%	**9%**
Severe challenging behavior	4%	3%	3%	3%	3%	**3%**

Table 3

Race/Ethnicity of GSRP Participants, 2017–2022

Year	2017–2018	2018–2019	2019–2020	2020–2021	2021–2022	Totals
Number of Students	38,088	38,075	37,232	28,422	36,415	**178,232**
White	54%	53%	54%	57%	52%	**54%**
Black	27%	28%	27%	24%	29%	**28%**
Hispanic	10%	10%	11%	11%	10%	**10%**
Multiracial	5%	5%	6%	5%	6%	**5%**
Asian	2%	2%	2%	2%	2%	**2%**
American Indian/ Alaska Native	<1%	1%	<1%	<1%	1%	**1%**
Hawaiian/Pacific Islander	<1%	<1%	<1%	<1%	<1%	**<1%**

Table 4

Distribution of Eligibility Factors by Class, 2017–2018 to 2021–2022 School Years

	Class 1 Higher income with selective risk	**Class 2** Low income with lower risk	**Class 3** Low income with average risk	**Class 4** Extreme low income with selective risk	**Class 5** Extreme low income with high risk
% of total	6.3%	23.8%	19.6%	43.9%	6.3%
Number of students	11301	42376	35008	78282	11265
Average FPL*	6.566	4.488	3.037	1.467	1.325
% with disability	22.7%	11.5%	10.3%	7.7%	22.0%
% with severe challenging behavior	4.7%	2.9%	3.2%	0.7%	11.6%
% with home language not English	4.5%	7.1%	11.3%	12.7%	1.2%
% with parents with low education attainment	6.3%	10.3%	15.6%	16.7%	34.0%
% with abuse/neglect	9.0%	7.1%	7.5%	1.3%	37.8%
% with environmental risk	34.7%	39.1%	46.6%	49.1%	85.8%

Note: Darker shading indicates higher risk in the case of FPL and larger proportions of children in that class for all other variables.

* FPL categories were recoded into an ordinal variable and treated as a continuous variable in the model estimation; values ranged from 1 (lowest income level) to 7 (highest income level), with 1 = 0% to 50% FPL; 2 = 51% to 100% FPL; 3 = 101% to 150% FPL; 4 = 151% to 200% FPL; 5 = 201% to 250% FPL; 6 = 251% to 300% FPL; 7 = 301% and above.

Figure 1

Class Prevalence, 2017–2018 to 2021–2022 School Years Combined

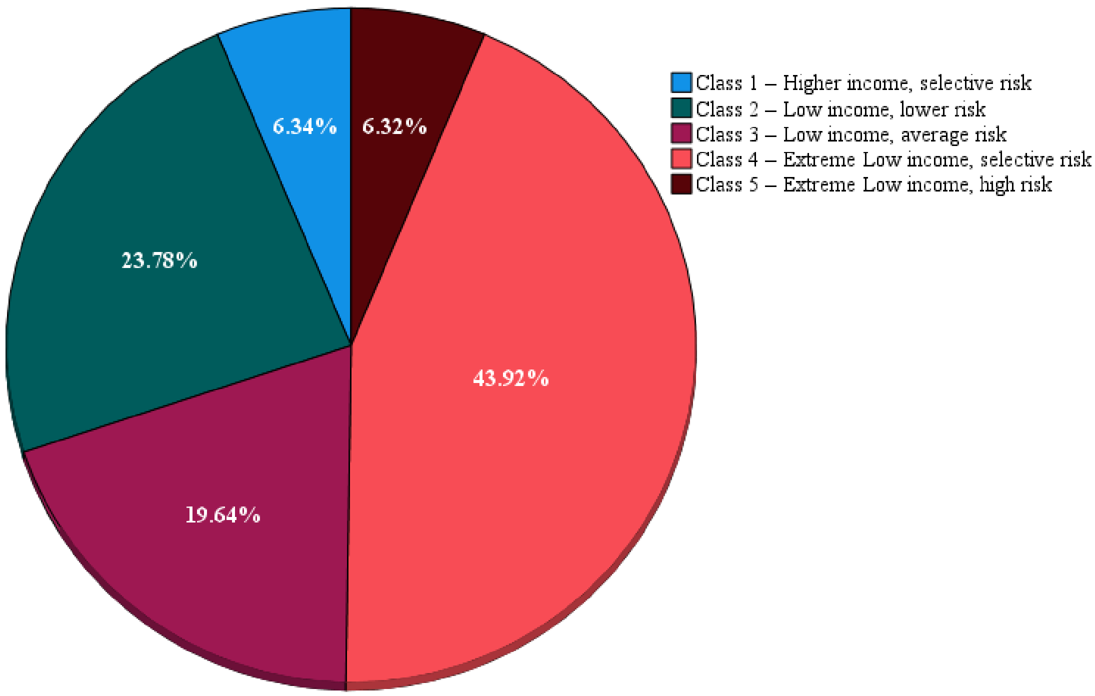


Figure 2

Conditional Risk of Eligibility Factors by Class, 2017–2018 to 2021–2022 School Years

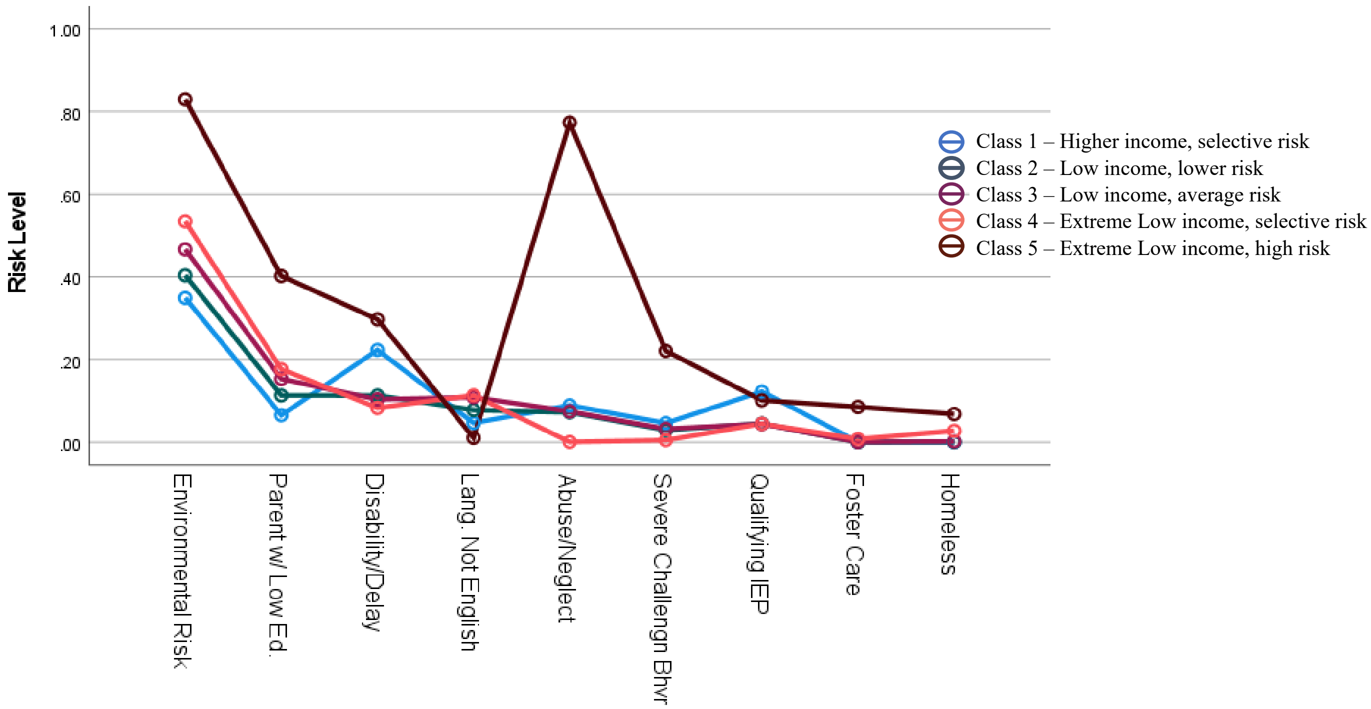


Figure 3

Prevalence of Classes by Year

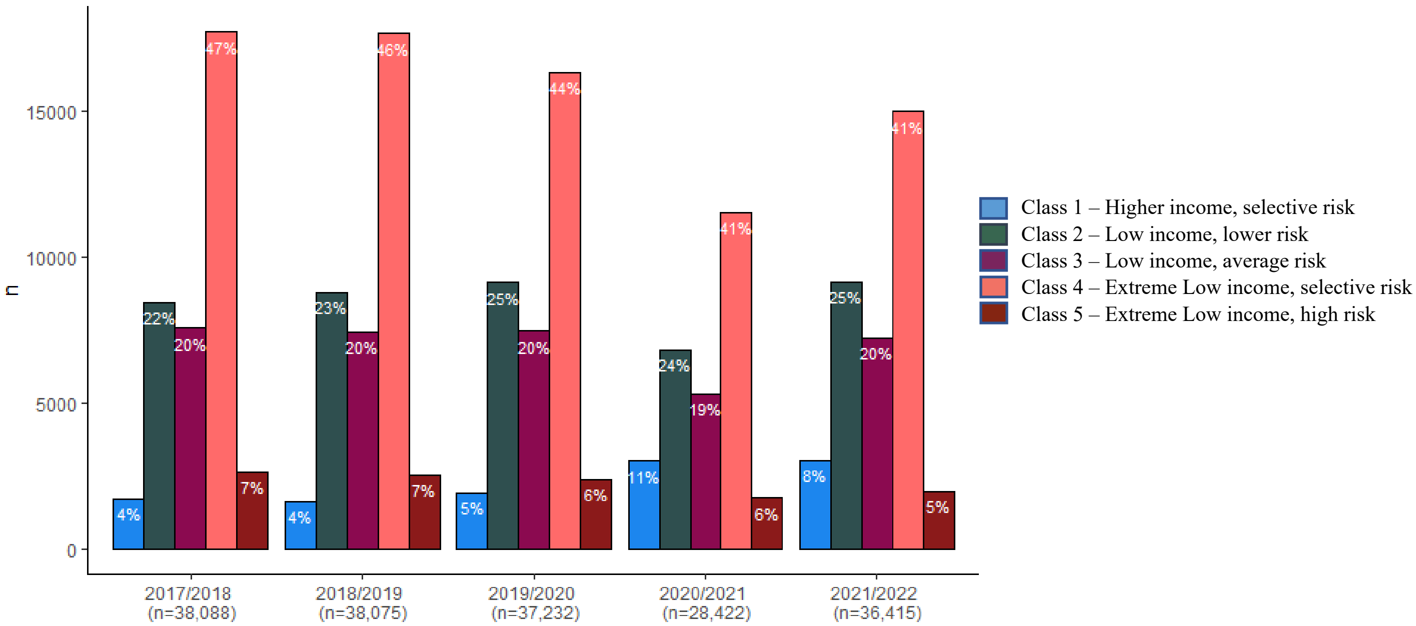
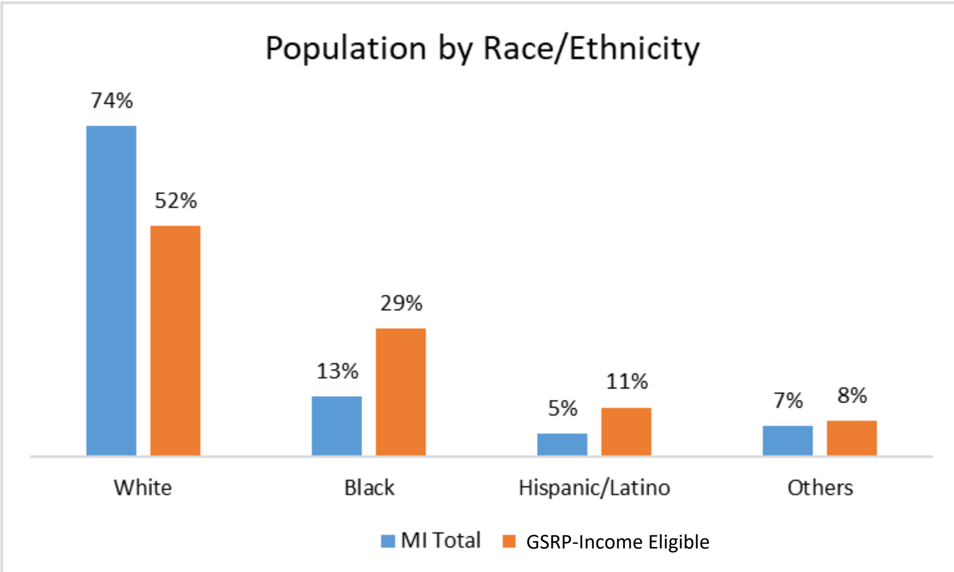


Figure 4

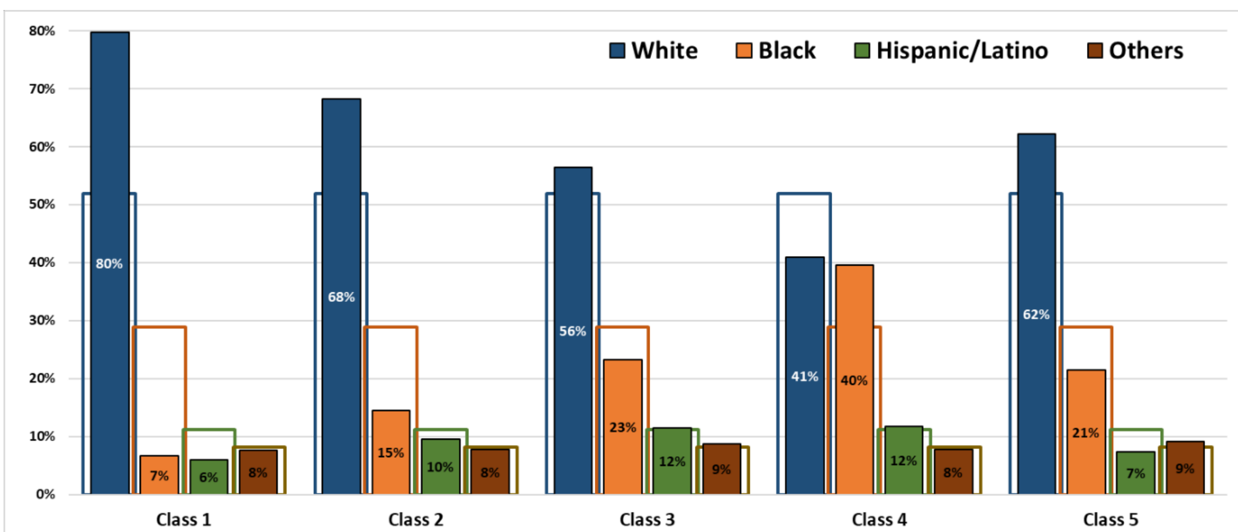
Michigan Total Population and GSRP Income-Eligible Population by Race/Ethnicity



Sources: For Michigan, U.S. Census Bureau American Community Survey 2020, whose five-year average covers roughly the study period of 2017–2021. For GSRP Income Eligible, Michigan public school records for kindergarteners in 2017–2021, showing families at 250% or less of FPL.

Figure 5

GSRP LCA Classes by Race/Ethnicity



Note: Transparent bars indicate the racial/ethnic makeup of the income-eligible population of Michigan children, derived from Michigan public school records for kindergarteners in 2017–2021.