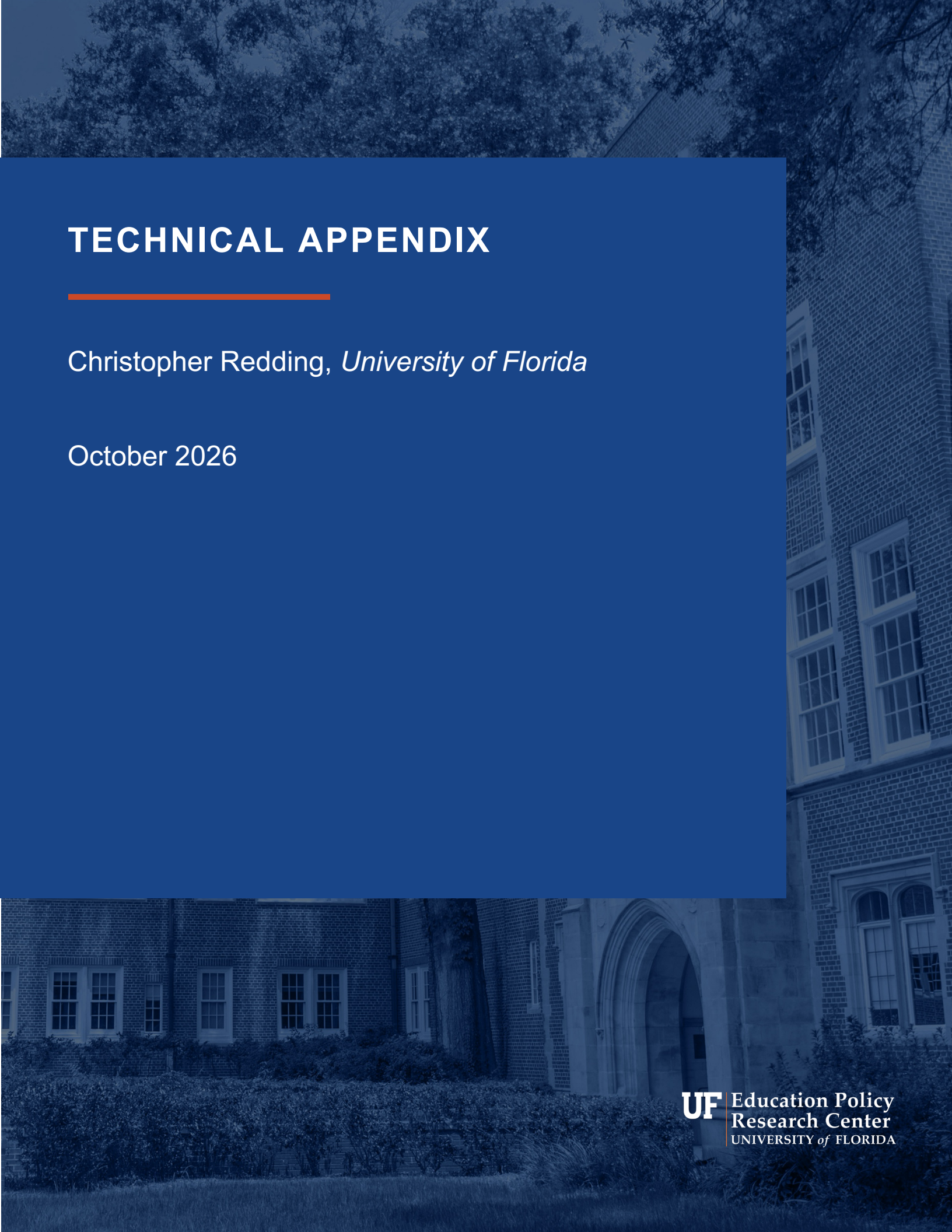


TECHNICAL APPENDIX

Christopher Redding, *University of Florida*

October 2026

UF Education Policy Research Center
UNIVERSITY of FLORIDA

UF Education Policy Research Center
UNIVERSITY of FLORIDA

TECHNICAL APPENDIX

University of Florida Education Policy Brief • October 2026

Christopher Redding

Introduction

This study draws on data from the Florida Department of Education (FLDOE) Staff Information System. These data include information on instructional and support staff working in K-12 schools in Florida. We draw primarily on annual data from the 2013-2014 to 2025-2026 school years that is collected in October of each school year (i.e., Survey 2). In some instances, Survey data is missing or incomplete for some school districts, in which case it is supplemented with Survey 3 data. Data are available for staff that are not exempt from Florida's broad public access to government records. These data were provided by the Florida Education Association, which requests these data from the FLDOE annually.

Identifying Unique Teacher Observations

A key challenge with using these data for research purposes is that they lack a unique teacher identification number from which to match teachers longitudinally. In its absence I develop an iterative matching process drawing primarily on teacher first and last names and date of birth. I first create a longitudinal dataset with all personnel files for each academic year. To construct a longitudinal panel dataset of teachers across years, it is critical to create a unique person identifier. To accomplish this goal, I used a combination of first name, last name, and date of birth, alongside other data (e.g., school district) for validation purposes. Prior to matching, first and last names were standardized by converting all text to uppercase, removing leading and trailing spaces, and collapsing multiple internal spaces. A preliminary teacher identifier was then created for all records sharing the same first name, last name, and date of birth.

Because administrative datasets frequently contain typographical errors and minor naming inconsistencies, I implemented several procedures to improve linkage quality. First, Soundex phonetic coding was used to identify potential misspellings of first and last names among records sharing the same date of birth and corresponding first or last name. When multiple spellings appeared to represent the same teacher, the unique identification variable was harmonized using the most commonly occurring identifier within the group. Second, I examined name changes over time by calculating edit distance (i.e., Levenshtein distance) between consecutive observations within the panel. For teachers with the same date of birth, changes of two or less characters were treated as likely spelling variations and the identification number was standardized accordingly. I then checked for potential change in last names across years among individuals with the same first and date of birth. This check was conducted for teachers working in the same district before being broadened statewide.

When I felt confident that unique identification numbers were created for all teachers in the sample, duplicate teacher-year observations were investigated. Teachers with missing school

assignments were removed when duplicate observations existed for the same teacher-year combination. Additional duplicate records resulting from identical teacher-year-school-job-salary combinations were dropped. For teachers appearing multiple times within a year, indicators were created to distinguish between teachers with multiple school assignments in the same year and multiple job roles within the same school. Records with zero duty days, an employment length listed as zero, or implausible salary values were removed when they appeared as duplicate assignments.

A final step in this data cleaning was to create a primary teaching assignment. In doing so, I first dropped seemingly superfluous observations for teachers who were observed more than once in the same school year. This included dropping observations where the number of days on duty or their employment length was listed as zero. I then dropped duplicated observations if the same teacher was listed as working in the same school, in the same job, with the same salary in the same year. I similarly dropped teachers if there were duplicated observations of teachers in the same school in the same role and they had no salary listed for one of the observations. After these data cleaning steps less than 1% of teachers have multiple observations in the same year.

Unless otherwise noted, the sample is restricted to full-time, regular instructional staff. An average of 156,267 unique teachers are observed annually, which corresponds closely to the 156,787 reported by the National Center for Education Statistics.[1]

Teacher Characteristics

The Staff Information System includes data related to teacher demographic characteristics, credentials, job characteristics, and salary. In addition, we also use these data to measure teacher mobility.

Demographic characteristics include gender, race/ethnicity, and age. Gender is measured as a binary male/female variable. Race/ethnicity includes the following mutually exclusive racial groups: American Indian, Asian, Black, Hispanic, Multiracial, Pacific Islander, and White. Age is separated into the following categories in most analyses: less than 30 years old, 30 to 39, 40 to 49, 50 to 59, 60 to 69, and over 70.

Teacher credentials are measured as years of experience and highest degree. Teacher experience was derived from annual records of reported teacher experience and subjected to a series of data cleaning procedures to improve consistency in the panel dataset. First, observations in which reported experience dropped to zero after previously being reported were assumed to reflect data entry errors rather than true changes in experience. Missing values were then imputed using linear interpolation based on adjacent observations within the panel dataset. To ensure that experience accumulated in a logically consistent manner over time, interpolated values were adjusted so that experience could not decrease across years for the same teacher. Resulting values were rounded to the nearest whole year, and indicators were retained to identify observations for which experience had been interpolated or otherwise modified from the originally reported value. As there were still cases where teaching experience did not increase when regular, full-time teachers were observed in the panel dataset in

[1] National Center for Education Statistics, "Florida," *Digest of Education Statistics State Dashboard*, accessed July 2, 2026, https://nces.ed.gov/programs/digest-dashboard/state/florida.

subsequent years, I reconstructed teaching experience by using each teacher's initial reported experience as a baseline and adding one year for each observed year of regular full-time employment. Even after these changes there were still inconsistencies in reported experience, particularly in terms of teachers being reported as having zero years of experience for multiple years in a row. To avoid overcounting first year teachers, I reconstructed experience trajectories by identifying years of regular full-time employment and accumulating service at the teacher-year level. I then combined the reconstructed measure with the reported experience variable, retaining the maximum value in each year to correct cases where administrative records failed to reflect accrued teaching experience. It should be noted that, although these changes were done to better model teachers' experience trajectories, the new experience variable correlated very highly with the original variable (r = 0.97). Still, it should be emphasized that this approach may underestimate novice teachers in the state. Finally, teachers were classified into six experience categories (0–2, 3–5, 6–10, 11–20, 21–30, and 31 or more years of experience), with a separate category reserved for observations with missing experience information. Highest degree includes associates/child development associate (CDA), bachelor's, master's, Specialist, doctorate, or missing.

Teacher job characteristics include those related to teachers' primary teacher assignment and features of their employment. I use the FLDOE job assignment codes to categorize instructional staff into mutually exclusive instructional roles, including (1) prekindergarten, (2) self-contained elementary, (3) English language arts (ELA)/reading, (4) math and computer education, (5) natural sciences, (6) social studies, (7) world languages, (8) art, dance, drama, and music, (9) health, physical education, and athletic coaching, (10) exceptional student education, (11) career and technical education, (12) adult education, and (13) other instruction, which includes humanities, librarian or media specialist, ROTC, remedial education, dropout prevention, among other instructional assignments. Employment features included the standard number of days per year an employee in this job is scheduled to work, employment type (i.e., contracted, temporary, or regular), part-time or full-time, and the number of months scheduled to work. Finally, salary data is available, which, in analyses is reported in both nominal and inflation-adjusted dollars.

I also leverage the longitudinal nature of the data to measure teacher. This includes teacher attrition, which is measured as 1 if the teacher is not observed as working in a Florida public school in subsequent years and 0 otherwise. Switching to a non-instructional position is measured as 1 if the teacher is observed in an instructional position in one year and a non-instructional year in the next year of observed data and 0 otherwise. I also create binary variables for whether a teacher moved to a new school, whether within the same district or in another district in the state.

School Characteristics

The Staff Information System indicates the school(s) in which teachers worked, which enables me to merge in school characteristics from the Common Core of Data. School characteristics include urbanicity, school level (i.e., primary, middle, high, and other level), school type (i.e., regular school, special education school, vocational school, and alternative school), charter school, virtual school, enrollment, fraction free or reduced price lunch (FRPL) eligible students, and fraction of American Indian, Asian, Black, Hispanic, Multiracial, Pacific Islander, and White students. Throughout the briefs, FRPL status is equated with school poverty. I recognize that there are issues with this language (Domina et al., 2018). but use school poverty to increase legibility of the briefs

Data Integrity

Steps were also taken to ensure that data used for final analyses was sufficient to support key inferences. This included checking for outliers and extreme values, as well as monitoring trends in study variables over time, both statewide and by school district. State trends were remarkably stable over time. There were some fluctuations observed in terms of the number of teachers and the characteristics of teachers within some school districts. To identify such instances, I created a tag for districts that experienced a more than 30% year-over-year change in the number of teachers overall, or by teacher characteristics. Several school districts were identified as having large fluctuations in the number of teachers, particularly in the last few school years. This includes Broward (2023-2024 to 2025-2026 school years), Glades (2024-2025 school year), Hillsborough (2025-2026 school year), and Pasco (2023-2024 to 2025-2026 school years). These districts are maintained for most analyses as aggregate teacher characteristics still remain consistent within these districts. So as not to overstate teacher turnover levels, they are excluded from analyses for teacher mobility.

Other than these overall trends in the number of teachers, two districts had subject-specific data anomalies. In Columbia County, the number of ELA teachers drops from 177 in 2020 to 44 in 2021 (while the overall number of teachers remains constant). Though some teachers left, large numbers were recorded as teaching other subject areas, including elementary, health, social studies, and arts/music, among others. In Madison County, the number of elementary school teachers drops from 53 in 2022 to 19 in 2023, again, with a stable number of teachers in the district. It appears that this shift was due to a move towards departmentalized instruction as most teachers were now teaching either ELA or math.

Trends in teacher salaries were quite stable, with one exception. In the 2023-2024 school year, more than 90% of teachers in Escambia County are listed as having a very low salary (e.g., <$1000). As a result, this year's data is excluded from the analysis of teacher salaries.

School characteristics were also stable over time, with the exception of teachers working in charter schools, which was less consistent over time. Though the number and relative share of charter school teachers have increased from 2017 to 2025, this upward trend has not been consistent when looking year-to-year. For instance, the number of charter school teachers drops from 11,387 in 2018 to 10,008 in 2019 and then increases to 13,188 in 2020. This fluctuation mostly seems to result from increases or decreases in the numbers of charter school teachers in Miami-Dade and Duval counties.

Limitations

There are limitations with the current dataset that should be noted. First, there are some variables that are sometimes available in other teacher administrative datasets that are missing from the current dataset. This includes information about teacher licensure and entry pathway, as well as performance outcomes (e.g., observation ratings; evaluation scores). Second, though steps were taken to maintain data integrity, it is possible that variables used in this set of research briefs are measured with error given that they were collected for administrative rather than research purposes. Third, the dataset only includes information on the public school sector. Beyond not being able to measure the non-public teacher workforce in the state, it limits the usefulness of some variables used in this analysis. For instance, when teachers are observed leaving teaching in Florida public schools, they may simply be moving to the private sector. Fourth, I am unable to merge study data with student-level data. All demographic data is

reported at the school level. Last and more broadly, as with other administrative data, the study does not include information about teacher job attitudes that would be obtained from surveys (e.g., job satisfaction; retention intentions).

Appendix

Notes and Sources

Microsoft Copilot was used to assist with improving readability and identifying typographical and grammatical errors during the preparation of this technical appendix. All analyses, interpretations, findings, and conclusions were reviewed and verified by the authors, who assume full responsibility for the content.