

Department of Economics
Working Paper

The Great Recession and Public Education

Online Appendix

By

William N. Evans

Robert M. Schwab

Kathryn L. Wagner

Working Paper 2017-02

College of Business Administration



The Great Recession and Public Education

William N. Evans
Department of Economics
University of Notre Dame
Notre Dame, IN 46530
Wevans1@nd.edu
574-631-7039

Robert M. Schwab
Department of Economics
University of Maryland
College Park, MD 20742
schwab@econ.umd.edu
301-405-3487

Kathryn L. Wagner
Department of Economics
College of Business Admin.
Marquette University
Milwaukee, WI 53201
kathryn.l.wagner@marquette.edu
414-288-3409

Online Appendix

Constructing a Balanced Panel of School Districts

We have developed a balanced panel school district data set for this analysis. The key source of data here is the Common Core of Data (CCD) from the National Center for Education Statistics (NCES) which is available at the Local Education Agency (LEA) (district) level. The Universe Surveys in the Common Core are available annually for the 1986-87 to 2013-14 academic years and include data on enrollment, staff counts, and basic school demographics recorded in the Fall of the school year.

The Universe Surveys of the Common Core do not contain any financial information, however. We take data at the LEA level and merge the Universe Surveys to the Finance Survey (F-33) data from the Common Core. The F-33 data provides information on revenues, expenditures, and capital outlays for districts and is available annually for the 1991-92 to the 2012-13 school years. As we constructed the data set for our analysis, we found that the financial data in the F-33 contained some extremely large and small values. These values could be valid, but it is more likely that some districts incorrectly reported enrollments or expenditures. We therefore developed an algorithm to delete extreme values. First, we calculated the (unweighted) 95th and 5th percentile district in total per-pupil current expenditures for each state and year. We then deleted districts with

per-pupil expenditures greater than 150 per-cent of the 95th percentile of per-pupil revenues or less than 50 percent of the 5th percentile.

Along with eliminating extreme values, we further restrict our sample by only keeping school districts that had similar operating procedures and organization/level. Given that schools with different operating procedures and organizations often experience different costs (secondary schools are typically more expensive than elementary schools), we focus on districts that are unified across elementary and secondary schools. All schools in our sample are also regular districts.

The Universe Survey of the CCD contains staff counts broken down by position type. Certain positions, however, are more consistently reported than others (e.g. districts report the number of teachers, but not all keep track of the number of library support staff). To account for this, we create a balanced panel by excluding districts that failed to report information for multiple years and districts that were only in operation for a limited number of years during our study period.¹ We use five key variables to balance the districts in our dataset: teachers, guidance counselors, school administrators, school district administrators, and student enrollment. Teacher employment and student enrollment were the most consistently reported variables across districts. In order to stay in the balanced sample, a school district had to report these variables in at least 17 of our 20 year study period (SY 1993/94 to 2012/13). For the remaining staff variables – guidance counselors, school and school district administrators -- districts needed to report at least 12 years.

Once we exclude these districts, we then develop a procedure to fill in the missing information of the remaining districts in the balanced sample. Missing teacher employment and student enrollment are linearly interpolated using data reported by a district in the year before and the year after the missing year of data. Given this strategy, we eliminate all districts that failed to

¹ In some years, New York City Public Schools reported as 33 distinct geographic districts. We aggregated the components of these 33 districts to form one aggregate measure for New York City Public Schools in the dataset.

report teacher or student information in the first or last years of our study period. Missing information for guidance counselors, school administrators, and district administrators is generated by calculating the average ratios of these employment types to teachers by district over the entire study period. We then multiply this ratio by the teacher count for the corresponding missing year to estimate the number of employees. The final data set excludes many small districts. As a consequence, while the data set contains only about 71 percent of all regular districts, those districts account for 88 percent of all public school students.