Data and Programs

Learning Job Skills for Colleagues at Work: Evidence from a Field Experiment Using Teacher Performance Data

April 2019

This .zip contains three Stata program files:

```
epp-tables-figures.do
epp_est.ado
epp-prepare-data.do
```

The Stata .do file `epp-tables-figures.do` produces the results reported in the main paper and the online appendix.

`epp-tables-figures.do` frequently calls the .ado file `epp_est.ado`, which calculates treatment effect point estimates, as well as wild cluster bootstrap-*t* *p*-values and Fisher randomization test *p*-values.

The main data file used by `epp-tables-figures.do` is `epp-analysis.dta`, which is has "student-level" observations (more specifically, there is one row in the data for each unique commination of student, year, subject, and assigned teacher). Additionally, `epp-tables-figures.do` also uses "teacher-level" data on teacher classroom observation scores pre- and post-experiment, `epp-obsv-pre.dta`, and `epp-obsv-post.dta` respectively. Each of these three data files is created by the .do file `epp-prepare-data.do`.

`epp-prepare-data.do` itself takes as inputs several data files provided to the research team by the Tennessee Department of Education (TN-DOE). These files are described in the next paragraph. Under the terms of our agreement with TN-DOE and federal privacy laws (FERPA), we, the authors, cannot provide the data to any third party. Interested researchers must apply to TN-DOE.

The data files provided by TN-DOE and used by `epp-prepare-data.do` are: (1) Year-by-year data files with individual student test scores in math and reading/language arts. Data file names are of the form `student_ach_data_2014.dta`. (2) Year-by-year data files linking students to their assigned teacher(s) by subject. Data file names are of the form `teacher_link_2014.dta`. (3) Year-by-year data files with the scores from teachers' classroom observations. Scores are provided for each of the 19 skills for each observation. Data file names are of the form `obs_level_eval_data_2014.dta`. These three types of data are the main inputs. Additional

TN-DOE data files used by `epp-prepare-data.do` include pre-treatment test-score "value-added" scores for teachers in the file `tvaas_2013.dta`, and data on teachers' years of experience in `teacher_exp_masked_oct14.dta`.

`epp-prepare-data.do` also uses two additional data files: (1) `epp-treat-assign.dta`, which has one row for each of the 14 study schools with columns for treatment assignment and randomization block. And (2) `epp-matches.dta` which provides details of the teacher roles and pairings created by the matching algorithm. See the text and Appendix B for more details about the algorithm. The `epp-matches.dta` file has one row for each teacher who was assigned to a pair, both "target" and "partner" teachers with indicators for these two roles. The data file includes additional details for each "target" teacher: (a) For each of the 19 skills, an indicator $= 1$ if the target teacher was "weak" in the given skill at baseline, i.e., a score < 3. (b) The ID number for algorithm-assigned "partner" as well as observable characteristics of that partner teacher, e.g., pre-treatment "value-added" and years of experience. (c) For each of the 19 skills, an indicator $= 1$ if the assigned partner teacher was "strong" in the given skill at baseline, i.e., a score ≥ 4. (d) The "proportion skills matched." The denominator of this proportion is the number of skills where the target is "weak" (the indicator in (a) $= 1$). The numerator is the number of skills where the target is "weak" and the partner is "strong" (the indicator in (c) $= 1$). And (e) indicators for whether the target and partner taught the same grade and the same subject. As described in the text, the data in `epp-matches.dta` are for both treatment and control schools because the data were created by the matching algorithm before random assignment.