STRATEGIC DATA PROJECT

ANALYZE: HUMAN CAPITAL ANALYSIS GUIDE BETA

SDP TOOLKIT

FOR EFFECTIVE DATA USE IN EDUCATION AGENCIES

www.gse.harvard.edu/sdp/toolkit

Toolkit Documents

An Introduction to the SDP Toolkit for Effective Data Use

Identify: Data Specification Guide

Clean: Data Building Guide for Human Capital BETA

Connect: Data Linking Guide for Human Capital BETA

Analyze: Human Capital Analysis Guide BETA

Adopt: Coding Style Guide

SDP Stata Glossary

VERSION: 0.5

Last Modified: December 18th, 2013

4. **Analyze:** Human Capital Analysis Guide

Conduct analyses that help answer key questions in your agency.

Analyze: Human Capital Analysis Guide is a set of step-by-step instructions to help any analyst in an education agency generate data visualizations about teacher recruitment, placement, development, evaluation, and retention. Through **Analyze**, your previous work identifying, cleaning, and connecting data will generate actual analyses to inform decision-making in your agency!

HUMAN CAPITAL ANALYSIS GUIDE

By now, you should have identified, cleaned, and connected your data into two analysis files named `Student_Teacher_Year_Analysis` and `Teacher_Year_Analysis`. You will now use these final analysis files to generate a number of analyses along the teacher pathway of being recruited, placed, developed, evaluated, and retained.

Analyze Structure

With each analysis, you will find:

- a picture of the analysis, based on the synthetic data;
- **Purpose:** an explanation of each analysis' value, and its ability to support understanding of teacher career trajectories and effectiveness patterns in your agency;
- **Required analysis file variables:** the variables from the analysis file you will need;
- **Analysis-specific sample restrictions:** a list of restrictions that you will apply to define the sample for the analysis;
- **Ask yourself:** a set of questions to help interpret results and invite deeper inquiry;
- **Potential further analyses:** further analyses you may conduct to understand underlying causes or interventions needed (this section is included in some but not all analyses); and
- **Analytic technique:** how to produce the analysis step-by-step using your analysis file and code in Stata.

Analytic Samples

There are two main analytic samples: a **teacher sample** corresponding to the `Teacher_Year_Analysis` file and a **student sample** corresponding to the `Student_Teacher_Year_Analysis` file created in **Connect**. These samples are further restricted using the **analysis-specific sample restrictions** specific to the individual analyses using indicator variables, such as `t_is_teacher`, `t_novice`, `t_experience`, and others you created in **Connect**. In the next section, we describe the **teacher sample** and **student sample** in greater detail, which you should read carefully if you are using your agency's own data to run these analyses.

Teacher sample: `Teacher_Year_Analysis`

The file will be unique by teacher ID (`tid`) and `school_year`, and includes many additional variables. For the full list of variables, refer to **Connect**.

tid	school_year	school_name	t_male	t_experience	other variables...
985	2004	Jackson Elementary	1	1	...
985	2005	Jackson Elementary	1	2	...
985	2006	Jackson Elementary	1	3	...

The teacher sample includes agency staff who:

1. received an agency salary in a given school year,
2. had a job code of "Teacher" in the same school year, and
3. were tied to students as the teacher-of-record in course or roster records in the same school year.

Generally, the teacher sample is defined by the indicator variable, `t_is_teacher`, which was generated in **Staff Task 3**: Staff Degrees and Codes in **Clean** (though it should be noted that in this toolkit we primarily use the second criterion as our sample data does not include agency salary, and links to students in a class for a teacher are tackled separately in **Student Task 4**: Student Class Enrollment). Also, in the event that teacher-student link data are poor or unavailable in your agency, you may define the teacher sample based on just the first two criteria, particularly for analyses that do not rely on estimates of teacher effectiveness.

There are a few **exceptions** to the three decision rules defined above:

- In some agencies, teaching positions are filled by long-term substitutes or other staff with job codes other than "Teacher." If this is the case, in your recruitment analyses you should consider separately reporting the number of new hires with "Teacher" job codes and the number of new hires with nonteaching job codes when many staff with nonteaching job codes are linked to students in course or roster records.
- When using a measure of teacher effectiveness, include all staff who serve as the primary teacher-of-record to students in tested grades and subjects. In your evaluation analyses, consider expanding the teacher sample to include all staff tied to students in tested grades and subjects, irrespective of their job codes. You may also consider excluding charter school teachers from your analysis if charter school data coverage is poor.
- When exploring patterns of teacher retention, you are examining the extent to which teachers remain in the classroom over time, as well as their patterns of transition into nonteaching jobs throughout their careers. For this reason, in your retention analyses restrict the analysis sample to include only teachers who meet the three decision rules in a given school year (Y), but ensure that those teachers in nonteaching positions in year (Y+1) are kept in the sample when reporting retention outcomes for year Y.
- As a specific example, say Jane is a teacher with a "Teacher" job code in 2006–07, and the following school year serves as an assistant principal. Two years later, however, she returns to teaching and has a "Teacher" job code once again. According to the decision rules, you should include Jane in the retention teacher sample in 2006–07. But in 2007–08, you will not include Jane in the teacher sample analyses because her job code in this school year is not "Teacher." You will, however, include her again in the retention Teacher Sample when she returns to teaching in 2008–09.
- One more caveat applies to retention analyses. First, teachers in the most recent school year for which data are available cannot be included, since we cannot identify the retention status of teachers when we have no way of observing their employment status in the following school year (because these data do not exist yet).
- Similarly, in the recruitment analyses we do not include teachers in the first year. Those data are not available, because new hire status is determined by whether or not a teacher was employed in the previous year.

Student sample: `Student_Teacher_Year_Analysis`

The file will be unique by student ID (`sid`), `school_year`, `tid_math` or `tid_ela`, and includes many additional variables. For the full list of variables, refer to **Connect**. To link students to teachers, this file is built upon a class-level file that lists the students taking a course with course and teacher information.

sid	school_year	tid_math	tid_ela	school_name	other variables...
1	2007	2657	.	Jackson Elementary	...
1	2007	.	2657	Jackson Elementary	...
2	2006	1354	.	Monroe Elementary	...
2	2006	.	5979	Monroe Elementary	...

In some analyses, you will be interested only in teachers and their employment patterns. In other analyses, however, you will need to rely on information about students, too. Our standard student sample includes all students in core courses (the result of Step 2 Restrict and Step 3 Generate in **Connect**) tied to the teachers in the teacher sample.

As a final note, it should be noted that source data for teachers and students often reflect the fact that students and teachers may transfer schools during the school year, and teachers may teach in more than one school. In these cases, use decision rules to simplify the structure of the data. For example, a good decision rule for assigning students to a single school is to include only students present in a school for more than half the school year.

Strategic Performance Indicators (SPIs)

Two analyses in the Placement and Retention sections of **Analyze** are part of the SDP Strategic Performance Indicators (SPIs), which provide deeper insight into the human capital performance of educational systems. These SPIs were produced using data from a number of SDP's partner agencies. You will be able to conduct a variant of these analyses yourself through **Analyze**.

You can read more about the SPIs at **www.gse.harvard.edu/sdp/spi**.

Running Your Analyses

At the beginning of your do file for each section of the human capital pathway, you will have a series of global switches indicating which analyses you would like to run. For example, here are the globals for the retention do file of the human capital pathway:

```
global four_year_trajectory              1
global four_year_trajectory_cert         1
global retention_by_VAM                  0
global four_year_trajectory_VAM          1
global retention_by_school_poverty       0
```

Each analysis is enclosed by a check for this global switch so that the code inside will run if the switch has a value of "1". In this way, you can control which analyses you want to run.

```
if $four_year_trajectory==1 {
...
```

}

Summary

After completing **Analyze**, you will have:

- used your final analysis files from **Connect** to generate and display many different analyses on teacher effectiveness,
- obtained new and confirmatory information about teachers in your agency, and
- learned essential methodologies to embark on your own "deeper dives" into the data.

Share these analyses with colleagues, peers, and senior leadership in your agency.

- Ask yourself how these analyses might further inquiry and inform policy.
- How might you adapt these analyses to track performance over time?
- What relationships were particularly informative?
- How might you extend certain analyses to be even more informative?
- Who should have this information?

As a final note, the analyses presented here do not capture all of our research team's efforts to understand human capital in education agencies. We believe the analyses presented are the most widely applicable to drive discussions about change. Moreover, we believe these analyses serve as a model to seek answers about human capital trends in your agency.

We would love to hear how these model analyses inspired different analyses and "deeper dives" in your agency. As always, if you require additional support, feel free to email us at **sdp@gse.harvard.edu**.

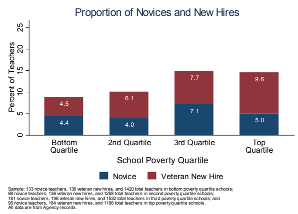
Map of Analyses

A. Recruitment

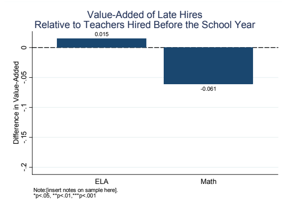
An examination of the kinds of teachers the agency hires, how many teachers it hires, what their preparation for teaching has been, and when they are hired.

New Teacher Hires by School Characteristics				
By School Poverty				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.441272966	0.413023679	0.465877712	0.337703199
Novice Hires	0.047244094	0.046903461	0.07495069	0.058206607
Veteran New Hires	0.892936803	0.886438809	0.83911939	0.827639752
By School Average Prior Math Score				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.322424893	0.2736	0.279844531	0.219581749
Novice Hires	0.070815451	0.0624	0.050527485	0.043250951
Veteran New Hires	0.780366057	0.771929825	0.819444444	0.803030303
By School Average Prior ELA Score				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.312682927	0.287042417	0.261011419	0.228613569
Novice Hires	0.070243902	0.064497385	0.05165851	0.039823009
Veteran New Hires	0.775351014	0.775303644	0.802083333	0.825806452

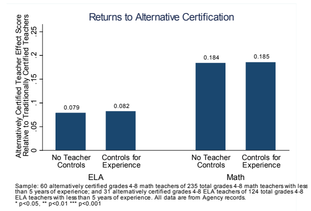
1. Table of Descriptive Information on Key Recruitment Practices
(p. 10)



2. New Hires by Poverty Quartile
(p. 13)



3. Value-added of Late Hires Relative to Teachers Hired Before the School Year Begins
(p. 16)



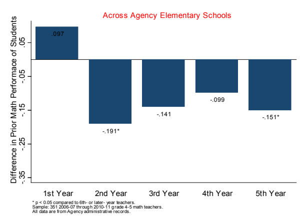
4. Value-added by Certification Pathway
(p. 21)

B. Placement

An examination of the patterns in student assignment to teachers across and within schools to identify places where efforts to reform placement policies could positively impact students and teachers.

Table of Teacher Characteristics by School Poverty Quartile				
Teacher Characteristics by School Poverty Level (2004-05 through 2011-12)				
	Low Poverty Schools	High Poverty Schools	Difference	N
Average Teacher Experience	13.652	11.537	-2.115**	15265
Novice Teacher	0.041	0.052	0.011**	15490
New Hire	0.086	0.114	0.028**	15490
Advanced Degree	0.498	0.427	-0.07**	16136
Alternative Certification	0.131	0.121	-0.01*	16661
National Board Certification	0.168	0.178	0.009	3861
Late Hire	0.062	0.068	0.006	1534
Previous 2-Year Pooled Math Teacher Effect	-0.001	0.003	0.004	557
Previous 2-Year Pooled English Teacher Effect	-0.005	-0.004	0.001	610

1. Table of Teacher Characteristics by School Poverty Quartile
(p. 26)

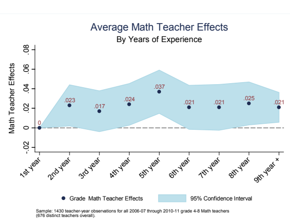


2. Prior Achievement of Students Placed with Teachers by Teacher Experience
(p. 29)

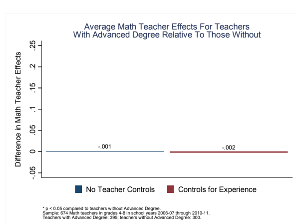
Strategic Performance Indicator

C. Development

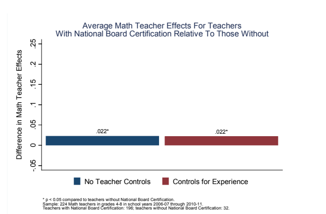
An examination of the ways teachers develop during their careers and an exploration of whether agency incentives are aligned with gains in teacher effectiveness.



1. Returns to Teaching Experience
(p. 36)



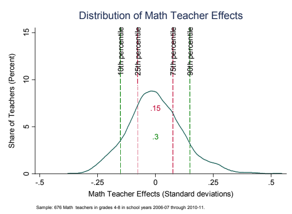
2. Returns to Advanced Degrees
(p. 40)



3. Returns to National Board Certification
(p. 44)

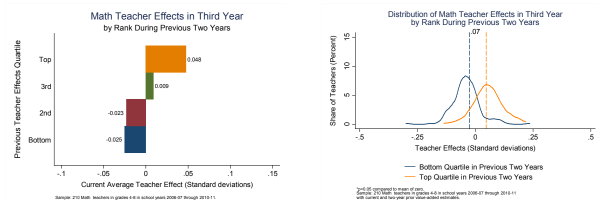
D. Evaluation

A good measure of teacher effectiveness will be spread out enough to distinguish exemplary teachers from developing ones in addition to being well correlated over time. The Evaluation section of the diagnostic examines the extent to which value-added estimates meet these criteria.



1. Distribution of Teachers by Value-Added Teacher Effect Estimates

(p. 51)

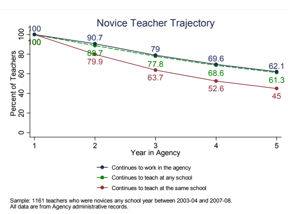


2. Predictive Power of Value-Added in Future Years Based on Prior Effectiveness Estimates

(p. 54)

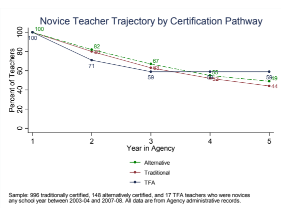
E. Retention

An examination of the types of teachers who transfer schools within the system, take nonteaching positions, and leave teaching in the agency altogether. This section examines how patterns vary across school characteristics, and among teachers with different teacher effectiveness estimates.



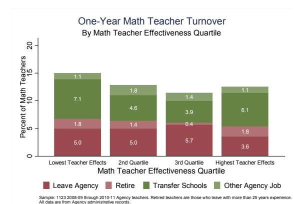
1. Four-Year Trajectory of Novice Teachers

(p. 60)



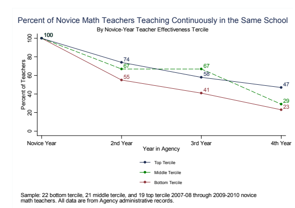
2. Four-Year Trajectory of Novice Teachers by Certification Pathway

(p. 64)



3. Retention by Teacher Effect Quartile

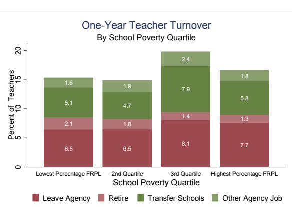
(p. 68)



4. Four-Year Trajectory of Novice Teachers by Teacher Effectiveness Tercile

(p. 71)

Strategic Performance Indicator



5. Retention by School Poverty Quartile

(p. 75)

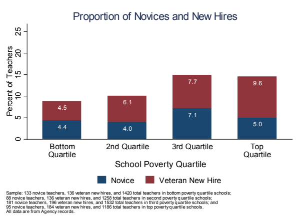
A. Recruitment

The recruitment process is the first opportunity education agencies have to secure highly effective teachers for their students. This section of the human capital pathway documents the kinds of teachers the agency hires (e.g., new hires and late hires), how many it hires, what their preparation for teaching has been, and when they are hired. Trends found here may be helpful in improving teacher recruitment and hiring by, for instance, identifying high-turnover schools, examining the demographic matches and mismatches between students and teachers, and determining which certification pathways produce the most effective teachers. Together, these trends can provide direction for a human resource strategy that aims to attract and place highly effective teachers in all classrooms.

New Teacher Hires by School Characteristics				
By School Poverty				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.441272966	0.413023679	0.465877712	0.337703199
Novice Hires	0.047244094	0.046903461	0.07495069	0.058206607
Veteran New Hires	0.892936803	0.886438809	0.83911939	0.827639752
By School Average Prior Math Score				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.322424893	0.2736	0.279844531	0.219581749
Novice Hires	0.070815451	0.0624	0.050527485	0.043250951
Veteran New Hires	0.780366057	0.771929825	0.819444444	0.803030303
By School Average Prior ELA Score				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.312682927	0.287042417	0.261011419	0.228613569
Novice Hires	0.070243902	0.064497385	0.05165851	0.039823009
Veteran New Hires	0.775351014	0.775303644	0.802083333	0.825806452

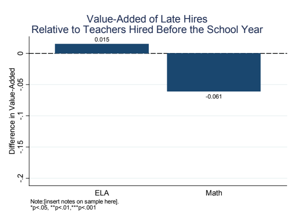
1. TABLE OF DESCRIPTIVE INFORMATION ON KEY RECRUITMENT PRACTICES

An agency snapshot of basic recruiting practices and the distribution of new hires.



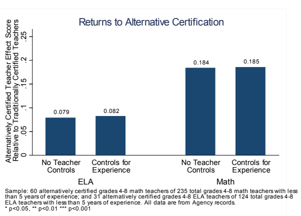
2. NEW HIRES BY SCHOOL POVERTY QUARTILE

Examines the extent to which new hires are distributed unevenly across the agency according to school characteristics.



3. VALUE-ADDED OF LATE HIRES RELATIVE TO TEACHERS HIRED BEFORE THE SCHOOL YEAR BEGINS

Determines if late hires are more or less effective at raising student achievement compared to teachers who are hired before the school year begins.



4. VALUE-ADDED BY CERTIFICATION PATHWAY

Determines if teacher effectiveness varies by certification status.

1. TABLE OF DESCRIPTIVE INFORMATION ON KEY RECRUITMENT PRACTICES

New Teacher Hires by School Characteristics				
By School Poverty				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.441272966	0.413023679	0.465877712	0.337703199
Novice Hires	0.047244094	0.046903461	0.07495069	0.058206607
Veteran New Hires	0.892936803	0.886438809	0.83911939	0.827639752
By School Average Prior Math Score				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.322424893	0.2736	0.279844531	0.219581749
Novice Hires	0.070815451	0.0624	0.050527485	0.043250951
Veteran New Hires	0.780366057	0.771929825	0.819444444	0.803030303
By School Average Prior ELA Score				
	Bottom Quartile	2nd Quartile	3rd Quartile	Top Quartile
New Hires	0.312682927	0.287042417	0.261011419	0.228613569
Novice Hires	0.070243902	0.064497385	0.05165851	0.039823009
Veteran New Hires	0.775351014	0.775303644	0.802083333	0.825806452

Purpose:

Obtain an agency snapshot of basic recruiting practices and the distribution of new hires.

Required analysis file variables:

```
school_year
t_newhire
t_is_teacher
t_novice
t_veteran_newhire
sch_pov_qrt
sch_avg_prior_math_qrt
sch_avg_prior_ela_qrt
```

other school characteristics of interest (agency-specific)

Analysis-specific sample restrictions:

- Keep only new hires whose job code is "Teacher."

Ask yourself:

- Are veteran and novice new hires distributed equitably and strategically across schools?
- Are there any trends over time in hiring novice vs. veteran new hires? For example, is the percentage of new hires who are novices increasing over time?

Potential further analyses:

Descriptive information on key recruitment practices disaggregated by teacher characteristics, such as race/ethnicity, content area, and preparation route.

1. TABLE OF DESCRIPTIVE INFORMATION ON KEY RECRUITMENT PRACTICES

Analytic technique: Calculate proportions, where the numerator is the number of new teachers in each school year, `t_newhire` (new teachers are those who do not appear in prior years of data). Novices or new hires are designated according to teacher experience data, and the denominator, `t_is_teacher`, is the total number of teachers in each school year.

```
/**** A. Recruitment ****/
/**** 1. Table of Descriptive Information on Key Recruitment Practices ****/

if $descriptive_table==1 {
```

// Step 1: Load the `Teacher_Year_Analysis` data file.

```
use "${analysis}\Teacher_Year_Analysis.dta", clear
```

// Step 2: Limit the sample to teachers.

```
keep if t_is_teacher==1
```

// Step 3: Create a text file to store the results.

```
file open myfile using "${graphs}\A1_Hires_by_School_Characteristics.xls", write replace
file write myfile "New Teacher Hires by School Characteristics" _n _n
```

// Step 4: Calculate the percentage of new hires, novices, and veteran new hires by school poverty quartile, and quartile of school average prior math and ELA scores.

```
foreach var of varlist school_poverty_quartile sch_avg_prior_math_qrt sch_avg_prior_ela_qrt {
```

// 1. Label each section of the table.

```
		if "`var'"=="school_poverty_quartile" {
			local title = "By School Poverty"
		}
		if "`var'"=="sch_avg_prior_math_qrt" {
			local title = "By School Average Prior Math Score"
		}
		if "`var'"=="sch_avg_prior_ela_qrt" {
			local title = "By School Average Prior ELA Score"
		}
```

// 2. Create column headings.

```
		file write myfile "`title'" _n
		file write myfile _tab "Bottom Quartile" _tab "2nd Quartile" _tab "3rd Quartile" _tab 
"Top Quartile" _n
```

1. TABLE OF DESCRIPTIVE INFORMATION ON KEY RECRUITMENT PRACTICES

// 3. Create row labels.

```
	foreach newhire of varlist t_newhire t_novice t_veteran_newhire {

		if "`newhire'"=="t_newhire" {
			local new = "New Hires"
		}
		if "`newhire'"=="t_novice" {
			local new = "Novice Hires"
		}
		if "`newhire'"=="t_veteran_newhire" {
			local new = "Veteran New Hires"
		}
```

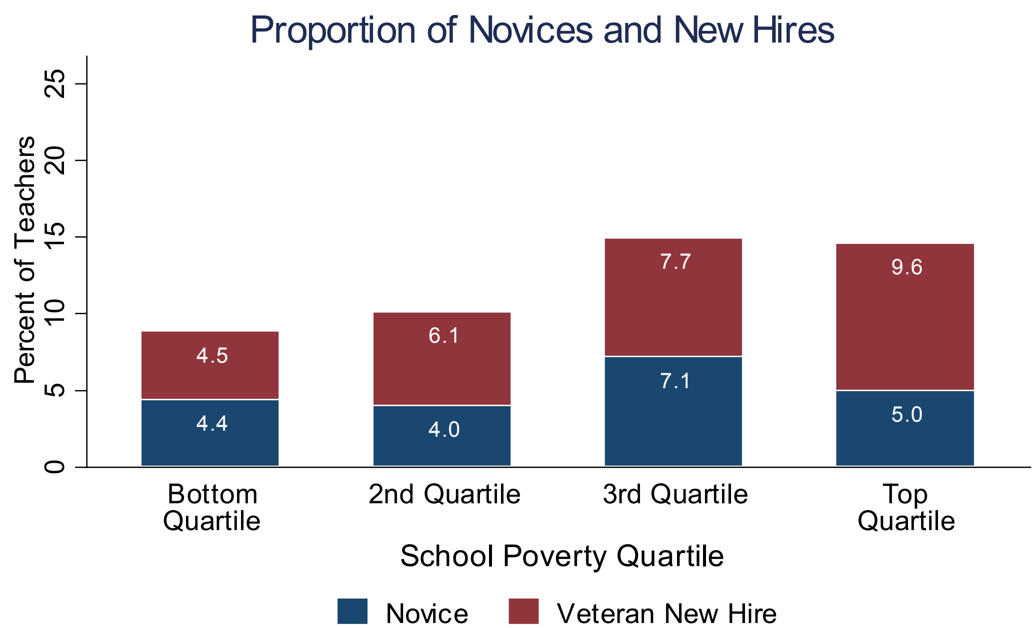
// 4. Calculate the mean for each cell.

```
		file write myfile "`new'" _tab
		forvalues q = 1/4 {
			sum `newhire' if `var'==`q'
			local quart_`q' = r(mean)
			di "`quart_`q'"
			file write myfile "`quart_`q''" _tab
		}
		file write myfile _n
	}

	file write myfile _n _n

}
```

// 5. Close the file. You can open and format the file in Excel.

```
file close myfile

}
```

2. NEW HIRES BY SCHOOL POVERTY QUARTILE



Sample: 133 novice teachers, 136 veteran new hires, and 1420 total teachers in bottom poverty quartile schools;
88 novice teachers, 136 veteran new hires, and 1258 total teachers in second poverty quartile schools;
181 novice teachers, 196 veteran new hires, and 1532 total teachers in third poverty quartile schools; and
95 novice teachers, 184 veteran new hires, and 1186 total teachers in top poverty quartile schools.
All data are from Agency records.

Purpose:
Examine the extent to which new hires are distributed unevenly across the agency according to school characteristics.

Required analysis file variables:

```
sch_pov_qrt
t_novice
t_veteran_newhire
t_is_teacher
```

Analysis-specific sample restrictions:

- Keep only employees whose job code is "teacher."

Ask yourself:

- How do hiring patterns differ between high- and low-poverty schools?
- Are the shares of novice and veteran new hires distributed equitably and strategically across school poverty quartiles?

Potential further analyses:
This graph is easily replicable to explore how the distribution of new hires varies across other school characteristics (e.g., AYP status, zone, school level, etc.).

2. NEW HIRES BY SCHOOL POVERTY QUARTILE

Analytic technique: Calculate the percentage of all teachers within each poverty quartile who are new hires. Among new hires, separately report those who are novices and those who enter the agency with prior teaching experience.

```
/**** A. Recruitment ****/
/**** 2. New Hires by School Poverty Quartile****/

if $school_poverty==1{
```

// Step 1: Load the `Teacher_Year_Analysis` data file.

```
use "${analysis}\Teacher_Year_Analysis.dta", clear
```

// Step 2: Calculate sample size.

```
forvalues qrt = 1/4 {
	unique tid if t_novice==1 & school_poverty_quartile==`qrt'
	local novice_`qrt' = r(sum)
	unique tid if t_veteran_newhire==1 & school_poverty_quartile==`qrt'
	local veteranhire_`qrt' = r(sum)
	unique tid if school_poverty_quartile==`qrt'
	local total_`qrt' = r(sum)
}
```

// Step 3: Collapse the teacher-level data file to include total counts of teachers and new hires within each poverty quartile.

```
collapse (sum) t_novice t_veteran_newhire (count) total_teacher=t_is_teacher, by(sch_pov_qrt)
```

// Step 4: Convert counts of new hires into a proportion (new hires/total teachers).

```
foreach v of varlist t_novice t_veteran_newhire {
gen pc_`v'=`v'/total_teacher*100
}
```

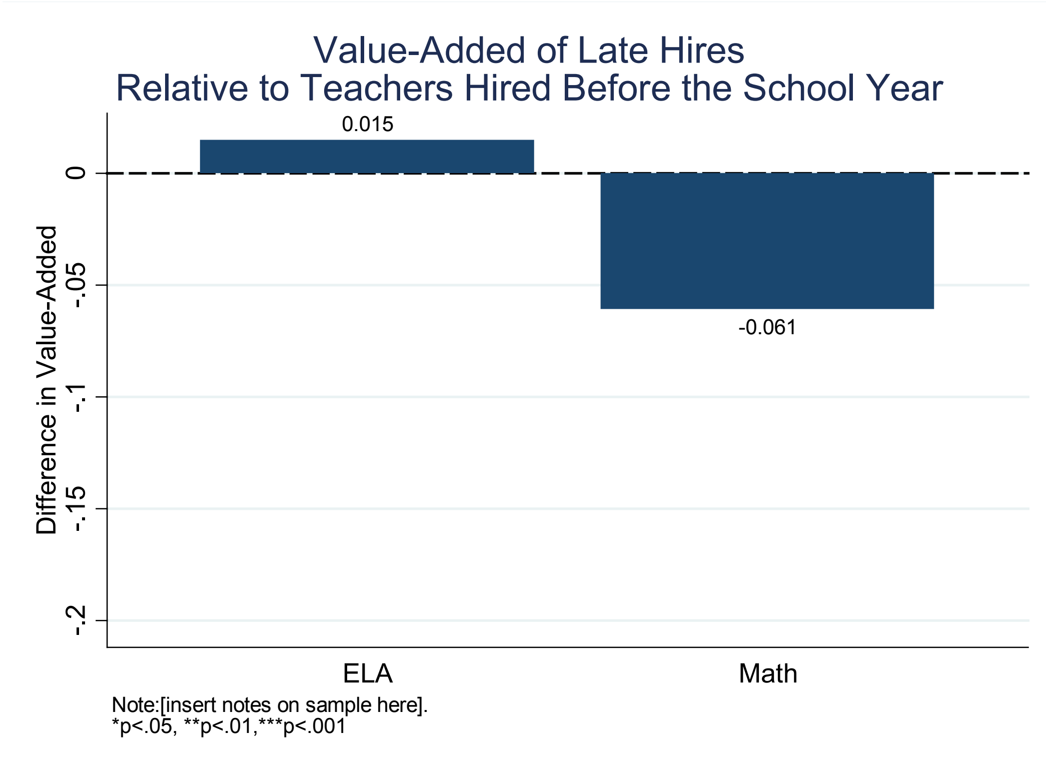
// Step 5: Create a bar graph of the outcomes.

```
#delimit ;
graph bar pc_t_novice pc_t_veteran_newhire,
over (sch_pov_qrt, relabel(1 `""Bottom" "Quartile""' 2 "2nd Quartile" 3 "3rd Quartile" 4 `""Top"
"Quartile""')) stack blabel(total, color(white) pos(inside)
format(%3.1f))
	ytitle("Percent of Teachers")
	title("Proportion of Novices and New Hires")
	b2title("School Poverty Quartile")
	yscale(range(0 25))
	ylabel(0(5)25, nogrid)
legend(region(lcolor(white)) symxsize(4) label(1 "Novice") label(2 "Veteran New Hire"))
	graphregion(color(white) fcolor(white) lcolor(white))
	plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `novice_1' novice teachers, `veteranhire_1' veteran new hires, and `total_1' total
```

2. NEW HIRES BY SCHOOL POVERTY QUARTILE

```
teachers in bottom poverty quartile schools;"
       "`novice_2' novice teachers, `veteranhire_2' veteran new hires, and `total_2' total teachers
in second poverty quartile schools;"
       "`novice_3' novice teachers, `veteranhire_3' veteran new hires, and `total_3' total teachers
in third poverty quartile schools; and"
       "`novice_4' novice teachers, `veteranhire_4' veteran new hires, and `total_4' total teachers
in top poverty quartile schools."
       "All data are from $agency_name records.", size(vsmall) span pos(7));
#delimit cr

graph export "${graphs}/A2_Hires_by_School_Poverty.emf", replace
graph save "${graphs}/A2_Hires_by_School_Poverty.gph", replace

}
```

3. VALUE-ADDED OF LATE HIRES RELATIVE TO TEACHERS HIRED BEFORE THE SCHOOL YEAR BEGINS



Purpose:
Determine whether late hires are more or less effective at raising student achievement as indicated by a measure of teacher effectiveness compared to teachers who are hired before the school year begins.

Required analysis file variables:

```
tid_math
tid_ela
school_year
t_experience
t_newhire
t_latehire
grade_level
std_scaled_score_math
std_scaled_score_ela
std_scaled_score_math_tm1
std_scaled_score_ela_tm1
cid_math
cid_ela
```

(see full list of variables in **Connect**):
student characteristics vector
class characteristics vector
cohort characteristics vector

Analysis-specific sample restrictions:

- Restrict to teachers with five or fewer years of teaching experience.
- Restrict to grades and subjects included in value-added estimates.
- Employ all other value-added estimate sample restrictions (see the value-added model technical appendix for the full list of sample restrictions).

Ask yourself:

- How are teacher effect estimates different for late and on-time hires?
- If late hires have lower teacher effect estimates, what supports might the agency give late hires to improve their performance?
- Given these results, what recruitment policies might the agency consider implementing or changing?
- If persistent differences in effectiveness are observed, to what extent is this differentially impacting certain types of schools? This question can be explored by examining the distribution of late hires across schools.

3. VALUE-ADDED OF LATE HIRES RELATIVE TO TEACHERS HIRED BEFORE THE SCHOOL YEAR BEGINS

Analytic technique: Regression specification with student, peer, and cohort controls, and grade-by-year fixed effects.

```
/**** A. Recruitment ****/
/**** 3. Value-Added of late hires relative to teachers hired before the school year begins ****/

if $VAM_late_hire==1 {

foreach subject in math ela {
```

// Step 1: Load the `Student_Teacher_Year_Analysis` data file.

```
use "${analysis}\Student_Teacher_Year_Analysis.dta", clear
```

// Step 2: Restrict the file to include only teachers with five or fewer years of experience, teachers who are new to the agency in one of the school years included in the analysis sample, and only teachers and students included in the sample for which teacher effects are estimated.

```
egen max_newhire=max(t_newhire), by(tid_`subject')
keep if max_newhire == 1
keep if t_experience <= 5
keep if !mi(tid_`subject')
```

// Step 3: Calculate the sample size.

```
unique tid_`subject' if t_latehire==1
local late_`subject' = r(sum)
unique tid_`subject'
local total_`subject' = r(sum)
```

// Step 4: Create a dummy indicator for teachers in the analysis file who were ever hired late.

```
egen ever_late_hire = max(t_latehire), by(tid_`subject')
```

// Step 5: Create dummy variables for each year of teaching experience.

```
tab t_experience, gen(exp)
rename exp1 first_year_teacher
rename exp2 second_year_teacher
rename exp3 third_year_teacher
rename exp4 fourth_year_teacher
rename exp5 fifth_year_teacher
```

// Step 6: Create interaction terms between the experience dummy variables and the ever hired late indicator.

```
foreach year in first second third fourth fifth {
       gen late_hire_`year'_year_teacher = ever_late_hire*`year'_year_teacher
}
```

// Step 7: Create grade-by-year fixed effects.

```
egen grade_by_year = group(grade_level school_year)
```

// Step 8: Estimate differences in teacher effectiveness between late hires and standard hires.

3. VALUE-ADDED OF LATE HIRES RELATIVE TO TEACHERS HIRED BEFORE THE SCHOOL YEAR BEGINS

// 1. Run the regression with interaction terms in the model.

```
areg std_scaled_score_`subject' ever_late_hire second_year_teacher third_year_
teacher fourth_year_teacher fifth_year_teacher late_hire_second_year_teacher
late_hire_third_year_teacher late_hire_fourth_year_teacher late_hire_
fifth_year_teacher std_scaled_score_`subject'_tm1 s_* _CL* _CO*,
cluster(cid_`subject') absorb(grade_by_year)
```

// 2. Test whether the interaction terms are significant in the model.

```
test _cons = 0
local ever_late_hire_sig = r(p)

foreach coef in second third fourth fifth {
	test _cons + late_hire_`coef'_year_teacher = 0
	local `coef'_sig = r(p)
}

gen sig_test = (`ever_late_hire_sig'>.05|`second_sig'>.05|`third_sig'>.05|`fourth_
sig'>.05|`fifth_sig'>.05)
sum sig_test
local sig_test_`subject' = r(mean)
drop sig_test
```

// 3. If the interaction terms in the model above are not statistically significant, omit the time in agency dummy variables and interaction terms from the regression model, and re-estimate the results including only the main effect of `ever_late_hire` controlling for teaching experience.

```
if `sig_test_`subject''==0 {
	areg std_scaled_score_`subject' ever_late_hire second_year_teacher third_year_teacher
fourth_year_teacher fifth_year_teacher std_scaled_score_`subject'_tm1, cluster(cid_`subject')
absorb(grade_by_year)
```

// 4. Store the math and ELA estimation results in a single data file for easy graphing.

```
	gen coef_ever_late_hire_`subject' =_b[ever_late_hire]
	keep coef*
	duplicates drop
	tempfile late`subject'
	save `late`subject''
}
else if `sig_test_`subject''==1 {
	gen coef_ever_late_hire_`subject' =_b[_cons]
	gen coef_ever_late_hire_yr2_`subject' = _b[_cons] + _b[late_hire_second_year_teacher]
	gen coef_ever_late_hire_yr3_`subject' = _b[_cons] + _b[late_hire_third_year_teacher]
	gen coef_ever_late_hire_yr4_`subject' = _b[_cons] + _b[late_hire_fourth_year_teacher]
	gen coef_ever_late_hire_yr5_`subject' = _b[_cons] + _b[late_hire_fifth_year_teacher]
	keep coef*
	duplicates drop
	tempfile late`subject'
	save `late`subject''
}
use `latemath', clear
append using `lateela'
gen time = 1
```

3. VALUE-ADDED OF LATE HIRES RELATIVE TO TEACHERS HIRED BEFORE THE SCHOOL YEAR BEGINS

```
	foreach v of varlist  coef_ever_late_hire_*  {
		egen max_`v'=max(`v')
		replace `v'=max_`v'
		drop max_`v'
	}
	duplicates drop
	reshape long coef_ever_late_hire_, i(time) j(subject) string
```

// Step 9: Graph the results.

// 1. If the interaction terms in the estimation model are not statistically significant, use the code below to produce the graph.

```
	if `sig_test_math'==0 | `sig_test_ela'==0 {
		#delimit ;
		graph bar coef_ever_late_hire,
			over(subj, relabel(1 "ELA" 2 "Math") gap(20))
			legend(cols(1) order(2 1)) ytitle("Difference in Value-Added")
			graphregion(color(white) fcolor(white) lcolor(white))
			plotregion(color(white) fcolor(white) lcolor(white))
			yscale(range(-.2 0.05)) ylabel(-.2(.05).05)
			blabel(bar, format(%9.3f) gap(*.6)) tit("Value-Added of Late Hires" "Relative to
Teachers Hired Before the School Year", span)
			note("Sample: `late_math' late grades 4-8 math teacher hires of `total_math'
total grades 4-8 new math teacher hires with less than"
			"five years of experience; and `late_ela' late grades 4-8 ELA teacher hires of
`total_ela' total grades 4-8 ELA new hires"
			"with less than five years of experience. All data are from $agency_name
records." "*p<.05, **p<.01,***p<.001", span pos(7))			yline(0, lpattern(dash)
lcolor(black));
		#delimit cr
		graph export "${graphs}/A3_Late_Hires_Value_Added.emf", replace
		graph save "${graphs}/A3_Late_Hires_Value_Added.gph", replace
	}
```

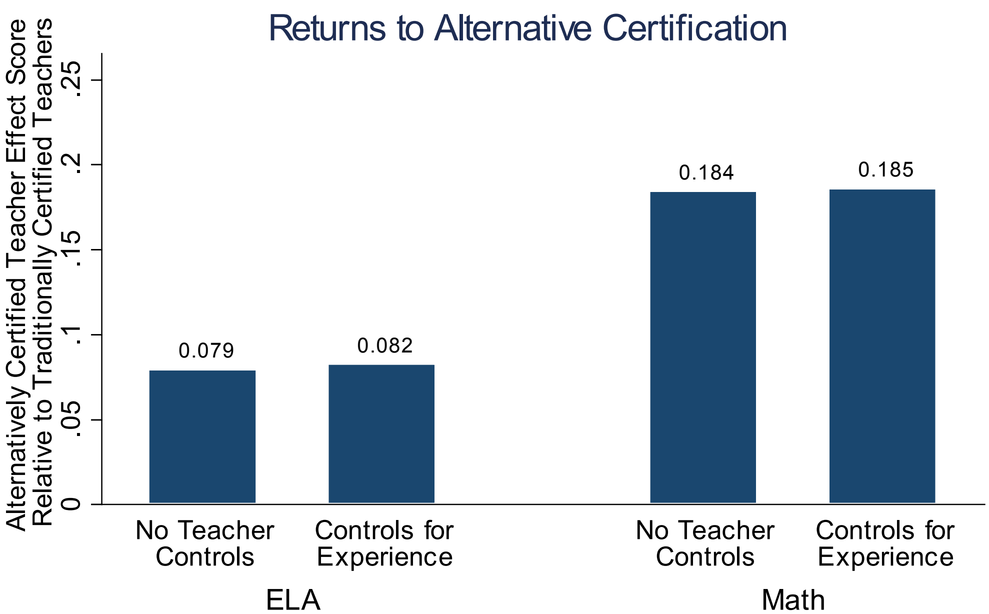
// 2. If the interaction terms in the estimation model are statistically significant, declare the dataset to consist of time series data, and use the code below to produce the graph.

```
	if `sig_test_math'==1 & `sig_test_ela'==1 {

		replace time = 2 if subject=="yr2_math" | subject=="yr2_ela"
		replace time = 3 if subject=="yr3_math" | subject=="yr3_ela"
		replace time = 4 if subject=="yr4_math" | subject=="yr4_ela"
		replace time = 5 if subject=="yr5_math" | subject=="yr5_ela"

		gen subj = subject
		replace subject = substr(subject, 5,.)
		replace subj = subject in 3/10
		drop subject

		reshape wide coef, i(time) j(subj) string
```

3. VALUE-ADDED OF LATE HIRES RELATIVE TO TEACHERS HIRED BEFORE THE SCHOOL YEAR BEGINS

```
             tsset time

             #delimit ;
                    tsline coef_ever*,  xlabel( 1 "Novice" 2 "2nd Year" 3 "3rd Year" 4 "4th Year" 5 
"5th Year" )
                    xtitle(" ") lwidth(thick) lwidth(thick) xscale(range(1 5.25)) yline(0, 
lcolor(black) lpa(dash)) legend(label(1 "Math") label(2 "ELA"))
                    tit("Value-Added of Late Hires" "Compared to Teachers Hired Before the School 
Year Begins", span)
                    subtit("By Years of Teacher Experience for Teachers Hired as Novice")
                    graphregion(color(white) lcolor(white)) plotregion(color(white) lcolor(white))
                    legend(region(lcolor(white)) symxsize(5) cols(1) pos(2))
                    note("Sample: `late_math' late grades 4-8 math teacher hires of `total_math' 
total grades 4-8 new math teacher hires with less than"
                    "five years of experience; and `late_ela' late grades 4-8 ELA teacher hires of 
`total_ela' total grades 4-8 ELA new hires"
                    "with less than five years of experience. All data are from $agency_name 
records." "*p<.05, **p<.01,***p<.001", span pos(7));
             #delimit cr

             graph export "${graphs}/A6_Late_Hires_Value_Added_Interactions.emf", replace
             graph save "${graphs}/A6_Late_Hires_Value_Added_Interactions.gph", replace
       }
}
}
```

4. VALUE-ADDED BY CERTIFICATION PATHWAY



Sample: 60 alternatively certified grades 4-8 math teachers of 235 total grades 4-8 math teachers with less than 5 years of experience; and 31 alternatively certified grades 4-8 ELA teachers of 124 total grades 4-8 ELA teachers with less than 5 years of experience. All data are from Agency records.
* p<0.05, ** p<0.01 *** p<0.001

Purpose:
Determine whether teacher effectiveness varies by certification status.

Required analysis file variables:

```
t_certification_pathway
current_tre_math
current_tre_ela
grade_level
school_year
std_scaled_score_math
std_scaled_score_ela
std_scaled_score_math_tm1
std_scaled_score_ela_tm1
cid_math
cid_ela
```

(see full list of variables in **Connect**):
student characteristics vector
class characteristics vector
cohort characteristics vector

Analysis-specific sample restrictions:

- Restrict to teachers with five or fewer years of teaching experience.
- Restrict to teachers with certification pathway information.
- Restrict to grades and subjects included in value-added estimates.
- Employ all other value-added estimate sample restrictions (see the value-added model technical appendix for the full list of sample restrictions).

Ask yourself:

- Are certain certification pathways associated with larger teacher effects? One way to explore this is by comparing teacher effect estimates across specific programs. Oftentimes, program-specific data are not available to perform this analysis, but when the data are accessible, we undertake such analyses (e.g., TFA vs. non-TFA teachers) using the analytic guidelines above.
- Given the results, what recruitment policies might the agency implement or change?
- What questions does this raise about teacher preparation and professional development in your agency?

Potential further analyses:
If you are a state, you might examine these trends across/by agencies to understand how location is linked to recruitment of teachers with alternative vs. traditional certification.

4. VALUE-ADDED BY CERTIFICATION PATHWAY

Analytic technique: Regression specification with student, peer, and cohort controls, and grade-by-year fixed effects.

```
/**** A. Recruitment ****/
/**** 4. Value-added by Certification Pathway ****/

if $VAM_certification_pathway==1 {
```

// Step 1: Load the `Student_Teacher_Year_Analysis` data file.

```
foreach subject in ela math {
	use "${analysis}\Student_Year_Analysis.dta", clear
```

// Step 2: Restrict the file to include only teachers with five or fewer years of experience, teachers who have certification pathway information, and only teachers and students included in the sample for which teacher effects are estimated.

```
	keep if t_experience <= 5
	keep if !mi(t_certification_pathway)
	keep if !mi(current_tre_`subject')
```

// Step 3: Create a dummy variable to indicate whether the teacher took an alternative certification pathway.

```
	gen alternative_certification = (t_certification_pathway>1 & t_certification_pathway!=.)
	tab alternative_certification t_certification_pathway, mi
```

// Step 4: Calculate sample size.

```
	unique tid_`subject' if alternative_certification==1
	local altcert_`subject' = r(sum)
	unique tid_`subject'
	local total_`subject' = r(sum)
```

// Step 5: Create dummy variables of teaching experience.

```
	tab t_experience, gen(exp)
	rename exp1 first_year_teacher
	rename exp2 second_year_teacher
	rename exp3 third_year_teacher
	rename exp4 fourth_year_teacher
	rename exp5 fifth_year_teacher
```

// Step 6: Create grade-by-year fixed effects.

```
	egen grade_by_year = group(grade_level school_year)
```

// Step 7: Estimate differences in teacher effectiveness between alternatively certified and traditionally certified teachers.

// 1. Estimate the results, including the main effect of `alternative_certification` controlling for teaching experience.

```
	areg std_scaled_score_`subject' alternative_certification second_year_teacher
	third_year_teacher fourth_year_teacher fifth_year_teacher std_scaled_
	score_`subject'_    tm1, cluster(cid_`subject') absorb(grade_by_year)

	estimates store m1
```

4. VALUE-ADDED BY CERTIFICATION PATHWAY

```
        // 2. Re-estimate the model above, without any controls for teacher experience.
        areg std_scaled_score_`subject' alternative_certification std_scaled_
        score_`subject'_    tm1, cluster(cid_`subject') absorb(grade_by_year)

        estimates store m2

        // 3. Store the math and ELA estimation results in a single data file for easy graphing.
        estimates restore m1
        gen coef_alt_cert_wexp_`subject' =_b[alternative_certification]
        keep coef*
        duplicates drop
        tempfile alt_wexp_`subject'
        save `alt_wexp_`subject''

        estimates restore m2
        gen coef_alt_cert_`subject' =_b[alternative_certification]
        keep coef*
        duplicates drop
        tempfile alt_`subject'
        save `alt_`subject''
}

use `alt_math', clear
append using `alt_wexp_math'
append using `alt_ela'
append using `alt_wexp_ela'
gen i = 1
foreach v of varlist  coef_alt_cert_*  {
        egen max_`v'=max(`v')
        replace `v'=max_`v'
        drop max_`v'
}

duplicates drop
reshape long coef_alt_cert_, i(i) j() string
rename _j subject
replace i = 2 if subject=="wexp_ela" | subject=="wexp_math"
ren i controls
replace subject = "ela" if subject=="wexp_ela"
replace subject = "math" if subject=="wexp_math"
```

4. VALUE-ADDED BY CERTIFICATION PATHWAY

// Step 8: Create a graph of the estimation results.

```
#delimit ;
	graph bar coef_alt_cert_, over(controls, relabel(1 `""No Teacher" "Controls""' 2
`""Controls for" "Experience""')) over(subject, relabel(1 "ELA" 2 "Math"))
bargap(10) blabel(bar, format(%6.3f)) legend( label(1 "No Teacher Controls")
	label(2 "Controls for Experience")) title("Returns to Alternative Certification")
yscale(range(0 .2)) ytick(0(.05).25) ytitle("Alternatively Certified Teacher Effect
Score" "Relative to Traditionally Certified Teachers")
	ylabel(0(.05).25, nogrid) graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
	note("Sample: `altcert_math' alternatively certified grades 4-8 math teachers of
`total_math' total grades 4-8 math teachers with less"
	"than 5 years of experience; and `altcert_ela' alternatively certified grades 4-8 ELA teachers
of `total_ela' total grades 4-8"
	"ELA teachers with less than 5 years of experience. All data are from $agency_name records."
"* p<0.05, ** p<0.01 *** p<0.001", span pos(7));

#delimit cr

graph export "${graphs}/A4_Alternative_Certification_Value_Added.emf", replace
graph save "${graphs}/A4_Alternative_Certification_Value_Added.gph", replace

}
```

B. Placement

Students are not randomly assigned to teachers' classrooms. Sometimes nonrandom assignment benefits students. For example, some teachers have a talent for working with hard-to-reach students and might get assigned to more low-performing students than other teachers. Unfortunately, some placement decisions have little to do with students' needs. When the most senior teachers concentrate in districts, schools, and classrooms with the most advantaged students, while novice teachers teach lower-performing students in hard-to-staff schools, achievement gaps widen. The SDP Placement analyses reveal patterns in student assignment to teachers across and within schools to identify places where efforts to reform placement policies could positively impact students and teachers.

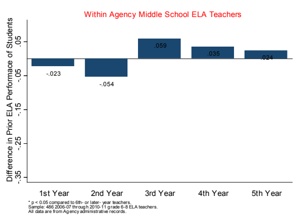
Table of Teacher Characteristics by School Poverty Quartile

Teacher Characteristics by School Poverty Level (2004-05 through 2011-12)

	Low Poverty Schools	High Poverty Schools	Difference	N
Average Teacher Experience	13.652	11.537	-2.115**	15265
Novice Teacher	0.041	0.052	0.011**	15490
New Hire	0.086	0.114	0.028**	15490
Advanced Degree	0.498	0.427	-0.07**	16136
Alternative Certification	0.131	0.121	-0.01*	16661
National Board Certification	0.168	0.178	0.009	3861
Late Hire	0.062	0.068	0.006	1534
Previous 2-Year Pooled Math Teacher Effect	-0.001	0.003	0.004	557
Previous 2-Year Pooled English Teacher Effect	-0.005	-0.004	0.001	610

1. TABLE OF TEACHER CHARACTERISTICS BY SCHOOL POVERTY QUARTILE

Examines the distribution of teachers across school characteristics.



2. PRIOR ACHIEVEMENT OF STUDENTS PLACED WITH TEACHERS BY TEACHER EXPERIENCE

Examines the placement relationship of students and teachers based on students' prior performance and teachers' experience.

1. TABLE OF TEACHER CHARACTERISTICS BY SCHOOL POVERTY QUARTILE

Table of Teacher Characteristics by School Poverty Quartile				
Teacher Characteristics by School Poverty Level (2004–05 through 2011–12)				
	Low-poverty Schools	High-poverty Schools	Difference	N
Average Teacher Experience	13.652	11.537	-2.115**	15265
Novice Teacher	0.041	0.052	0.011**	15490
New Hire	0.086	0.114	0.028**	15490
Advanced Degree	0.498	0.427	-0.07**	16136
Alternative Certification	0.131	0.121	-0.01*	16661
National Board Certification	0.168	0.178	0.009	3861
Late Hire	0.062	0.068	0.006	1534
Previous 2-Year Pooled Math Teacher Effect	-0.001	0.003	0.004	557
Previous 2-Year Pooled English Teacher Effect	-0.005	-0.004	0.001	610

Purpose:
Examine the distribution of teachers across school characteristics.

Required analysis file variables:

```
tid
school_code
school_year
t_is_teacher
t_latehire
t_novice
t_experience
t_newhire
t_adv_degree
certification_pathway
t_nbct
curr2year_tre_math
curr2year_tre_ela
school_poverty_quartile
```

Analysis-specific sample restrictions:
- Restrict the sample to teachers who are placed in schools in the top and bottom poverty quartiles.

Ask yourself:
- What supports could your agency offer to high-poverty schools with large shares of novice and early-career teachers?
- What screening tools do principals use across your agency? Does your agency provide training to principals about how to recruit and hire effective teachers?

Potential further analyses:
- Other teacher characteristics to add to the table if data are available include:
  - attended competitive postsecondary institution;
  - substitute teacher; and
  - master teacher, instructional mentor, other teacher leadership role.
- Explore to what extent these teacher characteristics are associated with value-added estimates of teacher effectiveness.

1. TABLE OF TEACHER CHARACTERISTICS BY SCHOOL POVERTY QUARTILE

Analytic technique: Calculate the averages of teacher characteristics across all schools within each school poverty category. Conduct a t-test to determine whether observed differences in teacher characteristics across school poverty categories are significant.

```
/**** B. Placement ****/
/**** 1. Table of Teacher Characteristics by School Poverty Quartile ****/

if $teacher_char_by_school_poverty==1 {
```

// Step 1: Load the `Teacher_Year_Analysis` data file.

```
use "${analysis}\Teacher_Year_Analysis.dta", clear
```

// Step 2: Merge data from the `School` file.

```
merge m:1 school_code using "${analysis}\School_Clean.dta", keep(1 3) nogen
```

// Step 3: Keep teacher and school characteristics of interest, and ensure that the data are structured to be unique by teacher and school year.

```
keep tid school_year school_code t_is_teacher t_latehire t_novice t_experience t_
newhire t_adv_degree certification_pathway t_nbct curr2year_tre_math curr2year_tre_
ela school_poverty_quartile

isid tid school_year
```

// Step 4: Keep only teachers working in low- or high-poverty schools in each school year.

```
keep if school_poverty_quartile == 1 | school_poverty_quartile == 4
```

// Step 5: Create an indicator to determine whether a teacher belongs to a school in the top poverty quartile. The others belong to the bottom poverty quartile.

```
gen sch_top_quart_pov = (school_poverty_quartile==4)
```

// Step 6: Create the table.

// 1. Generate binary variables for alternative certification.

```
gen alternative_certification = (certification_pathway>1 & certification_pathway!=.)
```

// 2. Define row titles in the table.

```
local t_experience "Average Teacher Experience"
local t_novice "Novice Teacher"
local t_newhire "New Hire"
local t_adv_degree "Advanced Degree"
local alternative_certification "Alternative Certification"
local t_nbct "National Board Certification"
local t_latehire "Late Hire"
local curr2year_tre_math "Previous 2-Year Pooled Math Teacher Effect"
local curr2year_tre_ela "Previous 2-Year Pooled English Teacher Effect"
```

// 3. Set up the Excel file for the table.

```
file open tbl using "${graphs}\B1_Placement_by_Poverty_Table.xlsx", write text replace
```

1. TABLE OF TEACHER CHARACTERISTICS BY SCHOOL POVERTY QUARTILE

```
        // 4. Create the table heading and column titles.
        file write tbl "Teacher Characteristics by School Poverty Level (2004-05 through 2011-12)"
        file write tbl _n
        file write tbl _tab " Low Poverty Schools"
        file write tbl _tab " High Poverty Schools"
        file write tbl _tab " Difference "
        file write tbl _tab _tab " N "
        file write tbl _n
```

// Step 7: For each teacher characteristic of interest, calculate the average difference between high- and low-poverty schools.

```
local varlist t_experience t_novice t_newhire t_adv_degree alternative_certification t_certification_
nbct t_latehire curr2year_tre_math curr2year_tre_ela
foreach tchr_char of local varlist {
        preserve
                reg `tchr_char' sch_top_quart_pov, robust
                estimates store `tchr_char'
                local lowpov = _b[_cons]
                local highpov = (_b[sch_top_quart_pov] + _b[_cons])
                local diff = _b[sch_top_quart_pov]
                test sch_top_quart_pov
                gen star =      ""
                replace star = "*" if  r(p) < .05
                replace star = "**" if r(p) < .01
                replace star = "**" if r(p) < .001

                display "`tchr_char'" _column(30) "`lowpov'" _column(60) "`highpov'" _column(90)
"`diff'" star
```

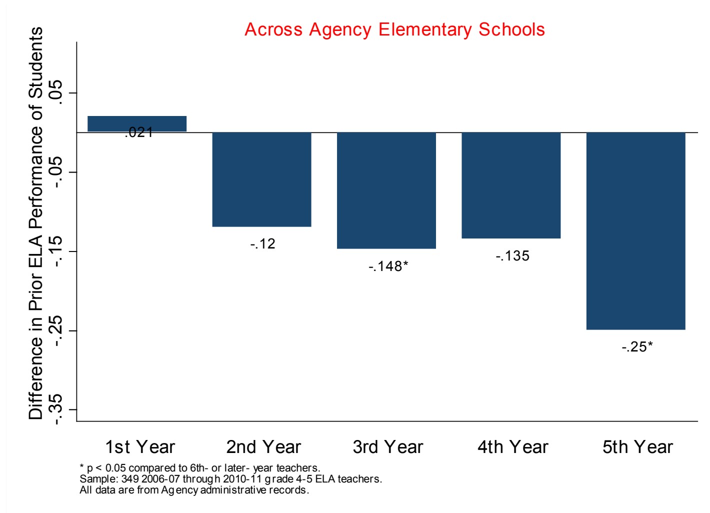
// Step 8: Capture the results from Step 5 in the table.

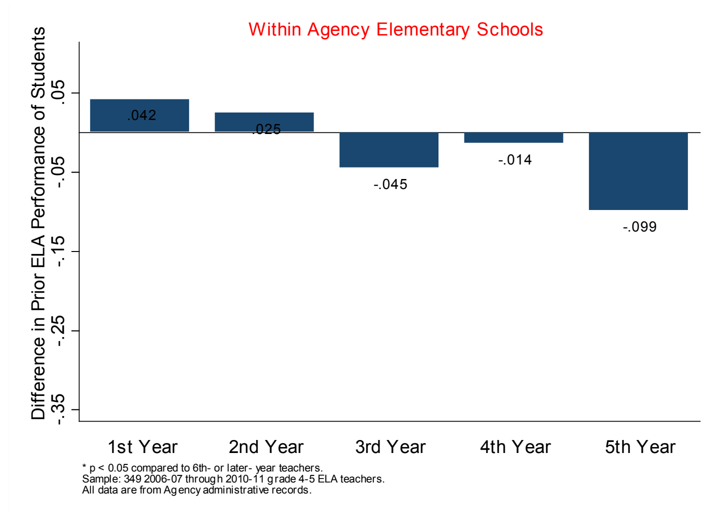
```
                file write tbl "``tchr_char''"
                file write tbl _tab "`:di %9.3f round(`lowpov', .001)'"
                file write tbl _tab "`:di %9.3f round(`highpov', .001)'"
                file write tbl _tab "`:di %9.3f round(`diff', .001)'"
                levelsof star
                file write tbl _tab "`:di %3s star'"
                file write tbl _tab "`:di %9.0f e(N)'"
                file write tbl _n

                drop star
        restore
}
file close _all
```

// Step 9: Format the table in Excel.

```
}
```

2. PRIOR ACHIEVEMENT OF STUDENTS PLACED WITH TEACHERS BY TEACHER EXPERIENCE





Purpose:

Examine how students are placed with teachers based on students' prior performance and teachers' experience.

Required analysis file variables:

```
tid_math
tid_ela
sid
school_year
t_experience
std_scaled_score_math
std_scaled_score_ela
std_scaled_score_math_tm1
std_scaled_score_ela_tm1
grade_level
cid_math
cid_ela
school_code
```

Analysis-specific sample restrictions:

- Restrict to classes included in value-added estimates.

Ask yourself:

- To what degree are the placement patterns driven by the concentration of novice teachers and lower-performing students in certain schools?
- Do internal school politics influence placement patterns? Are there formal or informal arrangements that enable more senior teachers to choose their classroom assignments? Is there a norm within the agency that novice teachers need to "put in their time" with more difficult assignments? Do parents of higher-achieving students influence placements to well-known teachers?
- Are timing factors important? Are classroom rosters drawn up early in the summer? Are students who enroll late assigned to teachers hired just prior to the school year?
- Are within-school gaps concentrated in certain schools? Are there some schools in which novice teachers are actually assigned to higher-achieving students?

Potential further analyses:

- This graph could be created for individual schools.
- If meaningful placement patterns are observed in middle school, examine how courses teachers are assigned (e.g., remedial, honors, etc.) vary based on teaching experience.

2. PRIOR ACHIEVEMENT OF STUDENTS PLACED WITH TEACHERS BY TEACHER EXPERIENCE

Analytic technique: Run a regression specification with grade-by-year fixed effects (and school fixed effects for within-school estimates).

- These placement results should be run separately for elementary and middle school teachers. Across the two estimates, ensure that grade-by-year dummies are mutually exclusive (this will prevent estimates that include highly singular matrices that result in the outputting of coefficients without estimated standard errors).
- Because teachers with six or more years of experience are the omitted group in the regression, the coefficients indicate average differences in prior performance of students assigned to teachers with five or fewer years of experience relative to teachers with six or more years of teaching history.

```
/**** B. Placement ****/
/*** 2. Prior achievement of students placed with teachers, by teacher experience ***/

if $student_prior_ach_by_tchr_exp==1 {

foreach subject in ela math {
```

// Step 1: Load the `Student_Teacher_Year_Analysis` data file.

```
	use "${analysis}\Student_Teacher_Year_Analysis.dta", clear
```

// Step 2: Keep only teacher, student, and school characteristics of interest for this analysis.

```
	keep tid_`subject' sid school_year t_experience std_scaled_score_`subject' std_
scaled_score_`subject'_tm1 grade_level cid_`subject' school_code
	duplicates drop
```

// Step 3: Create a variable denoting teacher experience in Years 1–5, as well as teachers with six or more years of teaching experience.

```
	local num = 1
	foreach v in first second third fourth fifth {
		gen `v'_year_tchr= (t_experience==`num')
		local num=`num'+1
	}
	gen sixthplus_year_tchr = (t_experience>5) & !mi(t_experience)
```

// Step 4: Create school-level indicators to allow for analysis to be conducted separately across elementary and middle schools if desired.

```
	gen elem=(grade_level < 6)
	gen middle=(grade_level > 5)
```

// Step 5: Generate grade-by-year fixed effects for inclusion in the regression model.

```
	egen grade_year=group(grade_level school_year)
	quietly tab grade_year, gen(_grade_year)
```

// Step 6: Generate school fixed effects for inclusion in the regression model.

```
	quietly: tab school_code, g(_s)
```

2. PRIOR ACHIEVEMENT OF STUDENTS PLACED WITH TEACHERS BY TEACHER EXPERIENCE

// Step 7: Estimate the assignment of students to teachers.

```
foreach span in elem middle {
       foreach model in across_schools within_schools within_teacher {
```

a. Set the model parameters.

```
if "`model'" == "across_schools"{
       local end = "if `span' == 1, absorb(grade_year)
cluster(cid_`subject')"
}
if "`model'" == "within_schools"{
       local end "_grade_year* if `span' == 1, absorb(school_code)
cluster(cid_`subject')"
}
if "`model'" == "within_teacher"{
       local end "_grade_year* _s* if `span' == 1, absorb(tid)
cluster(cid_`subject')"
}
```

b. Run the regression.

```
areg std_scaled_score_`subject'_tm1     first second third fourth fifth `end'
```

c. Store the estimates from the regression.

```
unique tid if e(sample) == 1
local n_`span'_`model' = r(sum)

foreach estimate in b se{
       gen `estimate'_`span'_`model'_yr1 = _`estimate'[first]
       gen `estimate'_`span'_`model'_yr2 = _`estimate'[second]
       gen `estimate'_`span'_`model'_yr3 = _`estimate'[third]
       gen `estimate'_`span'_`model'_yr4 = _`estimate'[fourth]
       gen `estimate'_`span'_`model'_yr5 = _`estimate'[fifth]
}
       } // end model loop
} // end grade span loop
```

// Step 8: Collapse the data to keep only one observation of each coefficient and its standard error for each variable in each regression.

```
collapse (max) b_* se_*
gen x = 1

#delimit ;
       reshape long
       b_elem_across_schools_yr se_elem_across_schools_yr
       b_elem_within_schools_yr se_elem_within_schools_yr
       b_elem_within_teacher_yr se_elem_within_teacher_yr
       b_middle_across_schools_yr se_middle_across_schools_yr
       b_middle_within_schools_yr se_middle_within_schools_yr
       b_middle_within_teacher_yr se_middle_within_teacher_yr
       , i(x) j(exp) ;
#delimit cr
```

2. PRIOR ACHIEVEMENT OF STUDENTS PLACED WITH TEACHERS BY TEACHER EXPERIENCE

// Step 9: Generate significance stars and prepare for graphing.

```
foreach span in elem middle {
       foreach model in across_schools within_schools within_teacher {
```

a. Set graph text.

```
if "`span'" == "elem"{
       local sch_type = "Elementary"
       local sch_type_ = "grade 4-5"
}
if "`span'" == "middle"{
       local sch_type = "Middle"
       local sch_type_ = "grade 6-8"
}
if "`model'" == "across_schools"{
       local subtitle = "Across ${agency_name} `sch_type' Schools"
       local model_save = "Across_Schools"
}
if "`model'" == "within_schools"{
       local subtitle = "Within ${agency_name} `sch_type' Schools"
       local model_save = "Within_Schools"
}
if "`model'" == "within_teacher"{
       local subtitle = "Within ${agency_name} `sch_type' School
`caps_subj' Teachers"
       local model_save = "Within_Teacher"
}
if "`subject'" == "math"{
       local caps_subj = "Math"
       local midsentence_subj = "math"
}
if "`subject'" == "ela"{
       local caps_subj = "ELA"
       local midsentence_subj = "ELA"
}
```

b. Generate significance stars.

```
gen sig_`span'_`model' = (abs(b_`span'_`model')-
                            (se_`span'_`model'*1.96)>=0)
replace b_`span'_`model' = round(b_`span'_`model',.001)
tostring sig_`span'_`model', replace
replace sig_`span'_`model' = "*" if sig_`span'_`model' == "1"
replace sig_`span'_`model' = ""  if sig_`span'_`model' == "0"
egen lab_`span'_`model' = concat(b_`span'_`model' sig_`span'_`model')
```

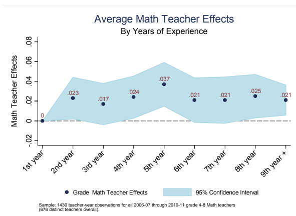
2. PRIOR ACHIEVEMENT OF STUDENTS PLACED WITH TEACHERS BY TEACHER EXPERIENCE

// Step 10: Graph the results.

```
		#delimit ;
			twoway (bar b_`span'_`model' exp,
			barwidth(.8) color(navy) fintensity(inten100))
			(scatter b_`span'_`model' exp,
			mlabel(lab_`span'_`model') msymbol(i)
			mlabpos(6) mlabcolor(black)),
			legend(off)
			title("`title_outcome'"
			 "For Early-Career `caps_subj' Teachers,"
			 "Compared to Experienced `caps_subj' Teachers")
			subtitle("`subtitle'", color(red) size(medsmall))
			ytitle("Difference in Prior `caps_subj' Performance of Students")
			xtitle("")
			xlabel(1 "1st Year" 2 "2nd Year" 3 "3rd Year"
			   4 "4th Year" 5 "5th Year", notick labgap(3))
			yscale(range(-.35 .10)) ylabel(-.35(.10).10, nogrid) ytick(-.35(.10).10)
yline(0,			lcolor(black) lwidth(vthin))
			graphregion(color(white) fcolor(white) lcolor(white)) plotregion(color(white)
	fcolor(white) lcolor(white))
			note("* p < 0.05 compared to 6th- or later- year teachers."
			 "Sample: `n_`span'_`model'' 2006-07 through 2010-11 `sch_type_' `midsentence_
subj'		teachers."
			 "All data are from ${agency_name} administrative records.", size(vsmall));

		#delimit cr

		graph save   "${graphs}\B2_Placement_by_Exp_`caps_subj'_`sch_type'_`model_save'.gph",
replace
		graph export "${graphs}\B2_Placement_by_Exp_`caps_subj'_`sch_type'_`model_save'.emf",
replace

		} // end model loop

	} // end school type loop

} // end subject loop

}
```

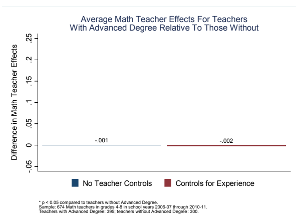
C. Development

Often, teachers are an agency's biggest investment. Once teachers have been recruited and placed in schools, their continued professional development benefits students and improves the success of the agency. Traditionally, agencies have incentivized two major forms of professional development: learning over time from experience and earning a graduate degree. This section of the diagnostic examines ways teachers develop during their careers and identifies whether agency incentives are aligned with gains in teacher effectiveness.



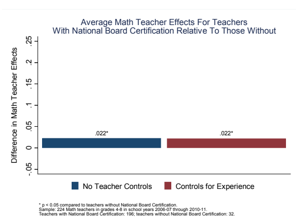
1. RETURNS TO TEACHING EXPERIENCE

Observes how teachers' value-added estimates change as they gain teaching experience.



2. RETURNS TO ADVANCED DEGREES

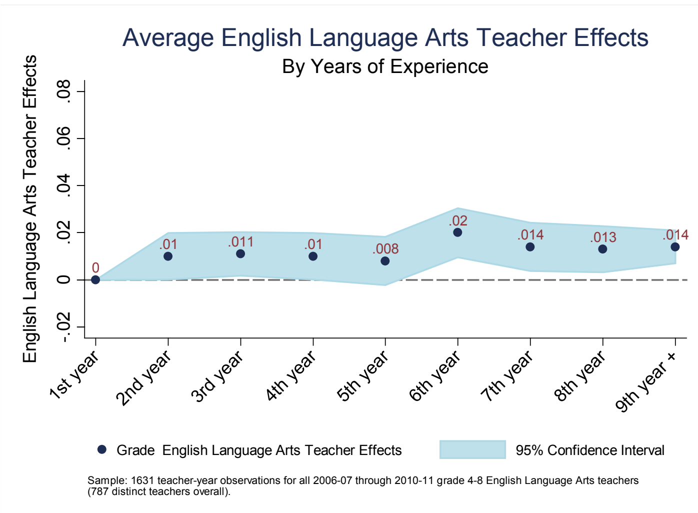
Determines if there are differences in value-added estimates between teachers with and without National Board Certification.

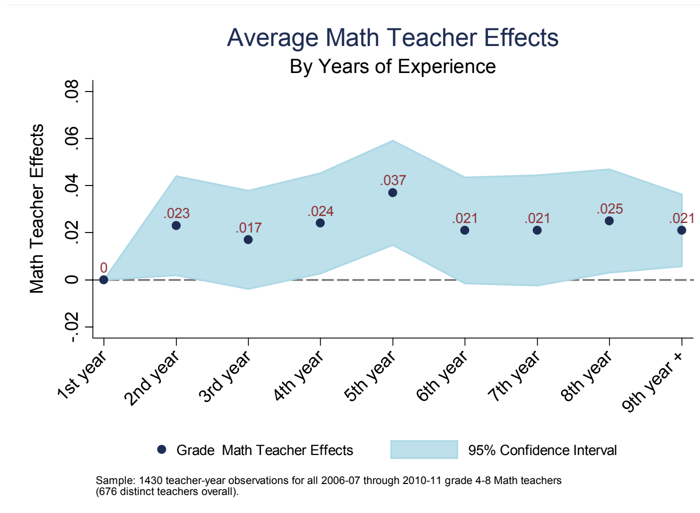


3. RETURNS TO NATIONAL BOARD CERTIFICATION

Determines if there are differences in value-added estimates between teachers who have National Board Certification and those who do not.

1. RETURNS TO TEACHING EXPERIENCE





Purpose:
Observe how teachers' value-added estimates change as they gain teaching experience.

Required analysis file variables:

```
tid
t_experience
current_tre_math
current_tre_ela
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers included in value-added estimates.

Ask yourself:

- Teacher salary schedules often compensate teachers for their teaching experience. How does your agency compensate teachers for experience? What are some ways that salary schedules could better align to increases in student outcomes over time?
- What induction and early career supports do novice teachers have? Do they vary by school, level of instruction (elementary, middle, high), and/or content area? Are the growth trajectories you see most related to recruitment practices, early career supports, or both?

Potential further analyses:

- If your agency changed induction and/or early career programs and/or policies, conduct this analysis separately for teachers who were and were not affected by the program or policy.
- Examine other dimensions of teacher effectiveness over time (e.g., trends in student survey results).

C. Development

1. RETURNS TO TEACHING EXPERIENCE

Analytic technique:

- Create eight binary variables for experience in Years 1–8, and a ninth indicator for experience in Years 9+. When displaying results, show only returns to experience in the first eight years of teaching.
- Regress the binary experience variables on the current year's teacher effect score separately for math and ELA. Optional: include student, peer, and cohort control variables and grade-by-year fixed effects.
- Illustrate the results with a confidence band. Make sure the standard errors from the regression estimates are multiplied by the appropriate number (e.g., 1.96) to create the desired confidence bands (e.g., 95%).

```
/**** C. Development ****/
/**** 1. Returns to Teaching Experience ****/

if ${return_to_experience} == 1 {

/*** Prepare the file for the analyses  ***/
{
```

// Step 1: Load the `Teacher_Year_Analysis` data file.

```
use "${analysis}/Teacher_Year_Analysis.dta", clear
```

// Step 2: Store long titles in a variable to be used in the graphs.

```
if "`subj'" == "math"{
       local subj_title "Math"
       local caps_subj "Math"
}
if "`subj'" == "ela"{
       local subj_title "English Language Arts"
       local caps_subj "ELA"
}
```

// Step 3: Create dummy variables of teaching experience.

```
forvalues x = 1/8 {
       gen exp`x' = (t_experience == `x')
}
gen exp9 = (t_experience > 8 & !mi(t_experience))
```

// Step 4: Save the prepared data into a temporary file.

```
tempfile tch_file
save `tch_file', replace
```

C. Development

1. RETURNS TO TEACHING EXPERIENCE

```
/*** 1. Create the average teacher effects by years of experience ***/

if ${return_to_experience} == 1 {
```

// Step 1: Load the teacher file.

```
use `tch_file', clear
```

// Step 2: Find average teacher effects for all experience categories relative to novice year, using regression.

```
reg current_tre_`subj' exp2-exp9
```

// Step 3: Gather all coefficients and standard errors.

```
forvalues x = 2/9 {
	gen est`x' =  _b[exp`x']
	gen se`x'  = _se[exp`x']
}
```

// Step 4: Get sample size by tagging observations that are included in the sample and counting the tagged observations.

```
gen _sample_ = e(sample)
unique tid if _sample_ == 1
local N_teacher = r(sum)
unique tid school_year if _sample_ == 1
local N_teacher_year = r(sum)
```

// Step 5: Collapse the data to capture the mean coefficients and standard errors for all experience categories.

```
collapse (mean) est* se*
```

// Step 6: Reshape to have an observation containing the coefficient and standard error for each experience level.

```
gen x = 1
reshape long est se, i(x) j(exp)
drop x
replace est = round(est,.001)
```

// Step 7: Create reference observation to indicate that the coefficient and standard error are zero for experience of one year.

```
local new_total_obs = _N + 1
set obs `new_total_obs'
replace exp = 1 if exp == .
replace est = 0 if exp == 1
replace  se = 0 if exp == 1
sort exp
```

// Step 8: Add the 5% confidence interval.

```
gen confidence_int_high = est + (se * 1.96)
gen confidence_int_low = est - (se * 1.96)
```

1. RETURNS TO TEACHING EXPERIENCE

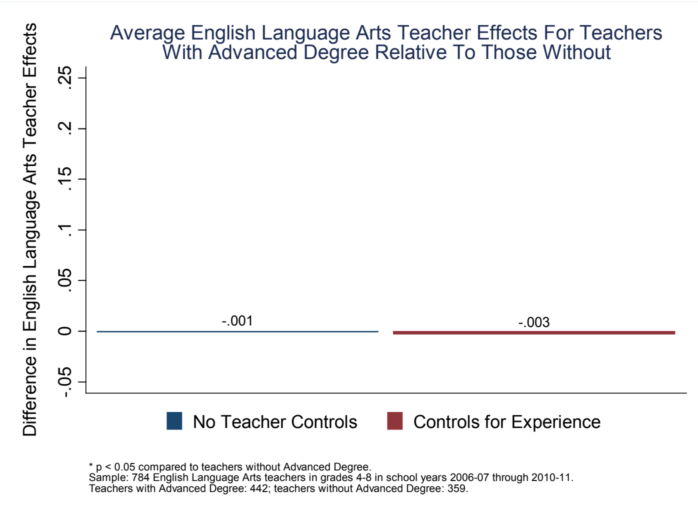
// Step 9: Generate a variable to indicate whether the estimate is significant.

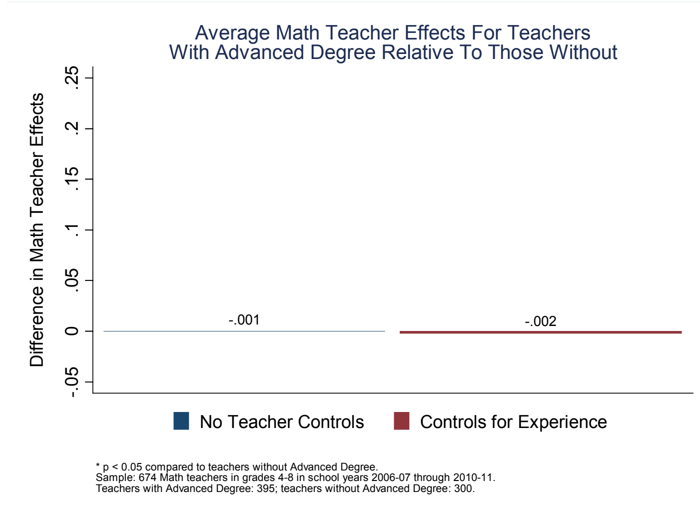
```
gen sig = (abs(est) - (se * 1.96) > 0)
```

// Step 10: Create the graph.

```
#delimit ;
	twoway (rarea confidence_int_high confidence_int_low exp, sort color(ltblue))
		(scatter est exp, mcolor(dknavy) mlabel(est) mlabpos(12)),
		yline(0, lcolor(gs7) lpattern(dash))
		xtitle("")
		ytitle("`subj_title' Teacher Effects", margin(0 3 0 0))
		legend(order(2 1) size(*.8) label(2 "Grade $sample_grade `subj_title'
	Teacher Effects") label(1 "95% Confidence Interval")
		region(lstyle(none) lcolor(none) color(none))) xtick(1(1)9)
		xlabel(1 "1st year" 2 "2nd year" 3 "3rd year" 4 "4th year" 5 "5th year"
	6 "6th year" 7 "7th year" 8 "8th year" 9 "9th year +", angle(45) notick
labgap(3))
		yscale(range(-.02 .08)) ytick(-.02(.02).08) ylabel(-.02(.02).08, nogrid)
		title("Average `subj_title' Teacher Effects")
		subtitle("By Years of Experience") ${graph_pref}
		note("Sample: `N_teacher_year' teacher-year observations for all 2006-
    07 through 2010-11 grade 4-8 `subj_title' teachers " "(`N_teacher'
distinct teachers overall).", size(vsmall));

#delimit cr

graph save "${graphs}\C1_Development_Return_to_Experience_`caps_subj'.gph" , replace
graph export "${graphs}\C1_Development_Return_to_Experience_`caps_subj'.emf" , replace

}
```

2. RETURNS TO ADVANCED DEGREES





Purpose:

Determine whether there are differences in value-added estimates between teachers with and without advanced degrees.

Required analysis file variables:

```
tid
t_adv_degree
current_tre_math
current_tre_ela
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers included in value-added estimates.
- If sample size allows, separate out PhDs from master's degrees. Otherwise, combine all forms of advanced degrees.

Ask yourself:

- How are teachers compensated for education beyond a bachelor's degree in your agency? Does your agency subsidize master's degrees as well as reward teachers who have them with higher salaries? What are some ways in which your agency can strengthen the link between compensation and actual performance?
- What alternatives can your agency offer to the master's degree bonus? How can your agency customize professional development for the individual teachers' areas for improvement, and reward teachers for progress toward their professional goals and student outcomes?

Potential further analyses:

- If degree title is available, separate teachers into groups by type of degree (e.g., administrative, content area that does or does not match subject taught).

2. RETURNS TO ADVANCED DEGREES

Analytic technique:

- Run a regression analysis to determine how teacher effectiveness varies for teachers with and without advanced degrees. Conduct the analysis separately for math and ELA teachers.
- Unlike the previous analysis—returns to experience—we do not estimate returns to an advanced degree (or National Board Certification) within teachers because teachers often work toward advanced degrees over several years. This information is seldom captured in the data. Including teacher fixed effects likely underestimates the returns to advanced degrees, although estimates are often small across teachers, too.

```
/**** C. Development ****/
/**** 2/3. Returns to Advanced Degrees/ Returns to National Board Certification ****/

/*** 2. Create the average teacher effects by teachers' advanced degrees ***/
if ${return_to_adv_degree} == 1 {
	local loop_list = "t_adv_degree"
	local graph_number = 2
	local var_graph = "Advanced Degree"
}

/*** 3. Create the average teacher effects by National Board Certification ***/
if ${return_to_natl_board} == 1 {
	local loop_list = "`loop_list' t_nbct"
	local graph_number = 3
	local var_graph = "National Board Certification"
}

foreach var in `loop_list' {
```

// Step 1: Load the teacher file.

```
	use `tch_file', clear
```

// Create a title for the graph, depending on the analysis.

```
		if "`var'" == "t_adv_degree"{
			local title_var = "Advanced Degree"
		}
		if "`var'" == "t_nbct"{
			local title_var = "National Board Certification"
		}
```

// Step 2: Find average teacher effects, without controlling for experience.

```
	reg current_tre_`subj' `var'
```

// Step 3: Gather coefficients and standard errors.

```
	gen est_base =  _b[`var']
	gen se_base  = _se[`var']
```

// Step 4: Get the sample size for prior estimate.

```
	unique tid if e(sample) == 1
	local teachers_in_sample_base = r(sum)
```

2. RETURNS TO ADVANCED DEGREES

// Step 5: Find average teacher effects, controlling for experience.

```
reg current_tre_`subj' `var' exp2-exp9
```

// Step 6: Gather coefficients and standard errors.

```
gen est_wexp =  _b[`var']
gen se_wexp  = _se[`var']
```

// Step 7: Get the sample size for prior estimate.

```
unique tid if e(sample) == 1
local teachers_in_sample_wexp = r(sum)
```

// Ensure that both samples have the same amount of teachers.

```
assert `teachers_in_sample_base' == `teachers_in_sample_wexp'
```

// Step 8: Get the number of teachers in the reference group.

```
unique tid if `var' == 0 & e(sample) == 1
local reference_group = r(sum)
```

// Step 9: Get the number of teachers in the treatment group.

```
unique tid if `var' == 1 & e(sample) == 1
local treatment_group = r(sum)
```

// Step 10: Collapse the data to capture the mean coefficients and standard errors for all experience categories.

```
collapse (mean) est* se*
```

// Step 11: Generate a variable to indicate the significance.

```
foreach spec in base wexp{
	gen sig_`spec' = ((abs(est_`spec') - 1.96 * se_`spec') > 0)
}
```

// Step 12: Reshape the data. This allows the use of the two-way bar instead of the graph bar.

```
gen x = 1
reshape long est_ se_ sig_, i(x) j(spec) string
replace spec = "1" if spec == "base"
replace spec = "2" if spec == "wexp"
destring spec, replace
```

// Step 13: Replace `sig_ dummy` with significance star to concatenate.

```
tostring sig_, replace
replace sig_ = "*" if sig_ == "1"
replace sig_ = ""  if sig_ == "0"
```

// Step 14: Create a variable that will serve as the bar labels.

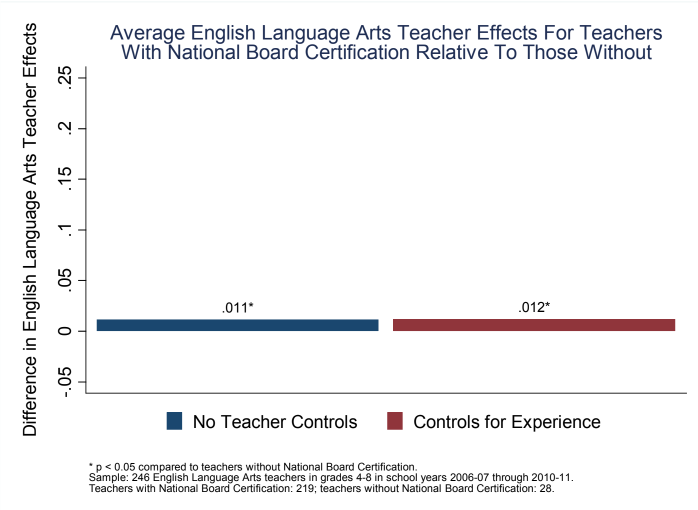
```
gen est_round = round(est_,.001)
egen est_label = concat(est_round sig_)
```

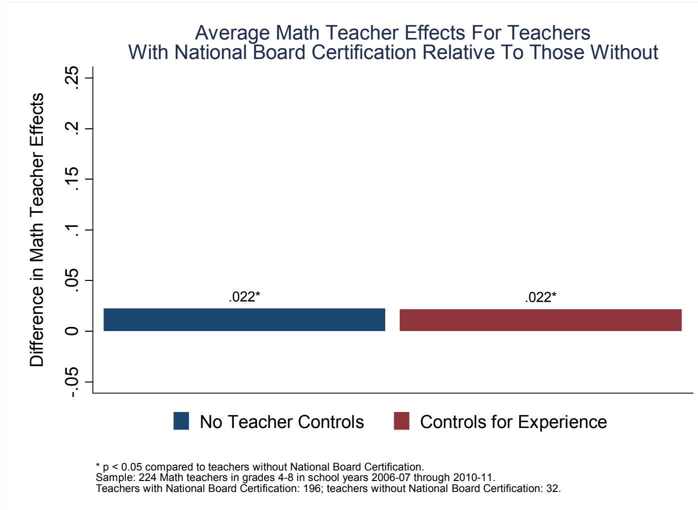
2. RETURNS TO ADVANCED DEGREES

// Step 15: Create the graph.

```
	#delimit ;
twoway		(bar est_ spec if spec == 1, barwidth(0.95) color(navy) fintensity(inten100))
		(bar est_ spec if spec == 2, barwidth(0.95) color(maroon) fintensity(inten100))
		(scatter est_ spec, mlabel(est_label) msymbol(i) mlabpos(12) 
	mlabcolor(black)),
		legend(label(1 "No Teacher Controls") label(2 "Controls for Experience") 
label(3 "") symxsize(3) ring(1) region(lstyle(none) lcolor(none) color(none)))
		title("Average `subj_title' Teacher Effects For Teachers" "With `title_var' 
	Relative	To Those Without", size(*.8))
		yscale(range(-.05 .25)) ytick(-.05(.05).25) ylabel(-.05(.05).25, nogrid)
		ytitle("Difference in `subj_title' Teacher Effects", margin(0 3 0 0))
		xtitle("") xlabel("", notick)
		${graph_pref}
		note("* p < 0.05 compared to teachers without `title_var'."
		"Sample: `teachers_in_sample_base' `subj_title' teachers in grades 4-8 in 
	school years 2006-07 through 2010-11."
		"Teachers with `title_var': `reference_group'; teachers without `title_var': 
	`treatment_group'.", size(vsmall)) ;

	#delimit cr

	graph save   "${graphs}\C`graph_number’_Development_Returns_to_`var_graph'_`caps_subj'.gph", 
replace
	graph export "${graphs}\C`graph_number’_Development_Returns_to_`var_graph'_`caps_subj'.emf", 
replace
```

3. RETURNS TO NATIONAL BOARD CERTIFICATION





Purpose: Determine whether there are differences in value-added estimates between teachers who have National Board Certification and those who do not.

Required analysis file variables:

```
tid
t_nbct
current_tre_math
current_tre_ela
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers included in value-added estimates.

Ask yourself:

- How does your agency support teachers who pursue National Board Certification? How are they rewarded for attaining the certification?
- How might your agency leverage the skills of National Board Certified teachers to improve teacher quality in their schools? Could National Board Certification be one path to a master teacher or teacher leader position in your district?

Potential further analyses:

- If there is available information, examine value-added by specific National Board Certification type. Also examine the match between certification type and teaching assignment.

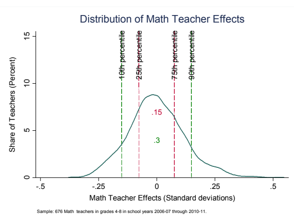
3. RETURNS TO NATIONAL BOARD CERTIFICATION

Analytic technique:

- Run a regression analysis to determine how teacher effectiveness varies for teachers with and without National Board Certification. Conduct the analysis separately for math and ELA teachers.
- Use the same code as the previous analysis, 2. RETURNS TO ADVANCED DEGREES, with your global `${return_to_natl_board}` set to 1.

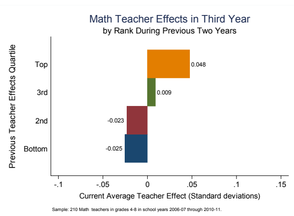
D. EVALUATION

Evaluating teachers serves two purposes: identifying areas where professional development is likely to benefit teachers, and identifying effective and ineffective teachers for career decisions. Evaluations can involve administrator observation, peer observation, collections of classroom preparation materials and artifacts, student surveys, and student achievement data. In the human capital diagnostic, we estimate teachers' effectiveness at raising student achievement using value-added methodology to evaluate teacher performance. A good measure of teacher effectiveness will have sufficient variation and consistency. That is, teacher effectiveness ratings are spread out across the range of possible values enough to observe differences across groups, and teachers' ratings are fairly well correlated over time. The Evaluation section of the diagnostic examines the extent to which value-added estimates meet these criteria.



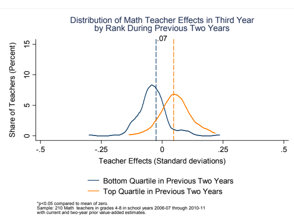
1. DISTRIBUTION OF TEACHERS BY VALUE-ADDED TEACHER EFFECT ESTIMATES

Examines the distribution of teacher effectiveness using value-added estimates.

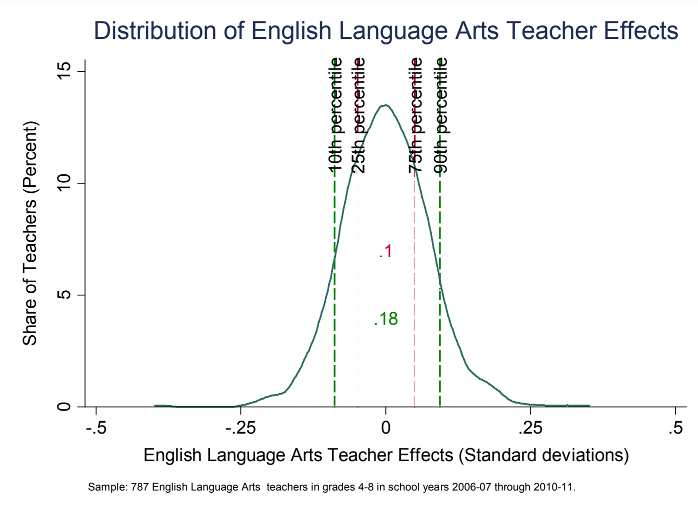


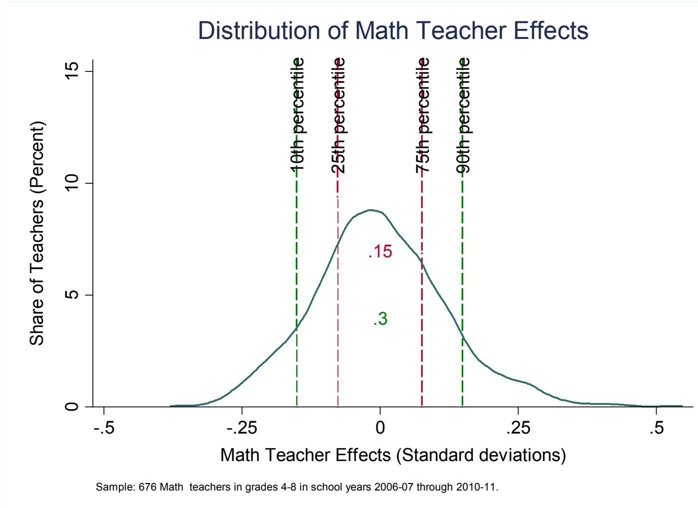
2. PREDICTIVE POWER OF VALUE-ADDED IN FUTURE YEARS BASED ON PRIOR EFFECTIVENESS ESTIMATES

Shows the extent to which prior value-added estimates of a teacher's effectiveness predict effectiveness in future years.



1. DISTRIBUTION OF TEACHERS BY VALUE-ADDED TEACHER EFFECT ESTIMATES





Purpose:
Examine the distribution of teacher effectiveness using value-added estimates.

Required analysis file variables:

```
tid
tre_math
tre_ela
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers included in value-added estimates.

Ask yourself:

- What measures of teacher effectiveness does your agency currently use? Compared to a measure that has little variation, what are the advantages of a measure of teacher effectiveness that has a lot of variation when making decisions about professional development, promotions to teacher leadership, and retention?
- What dimensions of teacher effectiveness do value-added estimates measure? What else is important to know about teacher quality? How are other dimensions of teacher quality measured in your agency?
- In what ways can knowledge of a teacher's value-added score be used to guide improvement? What kind of training would administrators and teachers need to be able to use the data to improve student achievement?

Potential further analyses:

- If more than one measure of teacher effectiveness is recorded in your agency, produce a kdensity graph for each one.
- Correlate two measures of teacher effectiveness (e.g., value-added estimates and classroom observation ratings). Create a scatterplot with the categories of performance on one axis and value-added ratings on the other. Examine the range of value-added estimates within each performance category.

1. DISTRIBUTION OF TEACHERS BY VALUE-ADDED TEACHER EFFECT ESTIMATES

Analytic technique: Produce a kdensity graph that depicts the distribution of value-added estimates. The distribution of teacher effects should be created based on all teachers' value-added estimates pooled together to take into account as many years of student achievement data as possible.

```
/**** D. Evaluation ****/
/**** 1. Distribution of Teachers by Value-Added Teacher Effect Estimates ****/

if ${overall_teacher_effects} == 1 {

/*** 1. Create the overall teacher effects analysis files and graphs ***/
```

// Loop through the subject areas (Math and ELA) to create the overall teacher evaluation analysis files.

```
foreach subj in math ela {
```

// Step 1: Load the `Teacher_Year_Analysis` data file containing value-added estimates.

```
	use "${analysis}/Teacher_Year_Analysis.dta", clear
```

// Step 2: Store long titles in a variable to be used in the graphs.

```
	if "`subj'" == "math"{
		local subj_title "Math"
		local caps_subj "Math"
	}
	if "`subj'" == "ela"{
		local subj_title "English Language Arts"
		local caps_subj "ELA"
	}
```

// Step 3: Restrict the sample to one observation per teacher.

```
	keep tid tre_`subj'
	keep if !missing(tre_`subj')
	duplicates drop
	isid tid

	count
	local samplesize = r(N)
```

// Step 4: Get percentiles, and store them in a local variable. This will enable you to position them on the graph.

```
	sum tre_`subj', detail
	local p10_`subj' = r(p10)
	local p25_`subj' = r(p25)
	local p75_`subj' = r(p75)
	local p90_`subj' = r(p90)
```

// Step 5: Get 10th–90th and 25th–75th percentile ranges, and store them in a local variable.

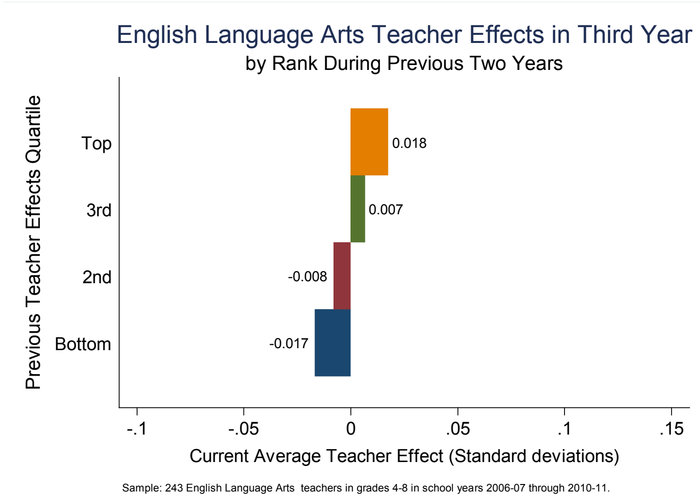
```
	local `subj'diff1090 = round(`p90_`subj'' - `p10_`subj'', .01)
	local `subj'diff2575 = round(`p75_`subj'' - `p25_`subj'', .01)
```

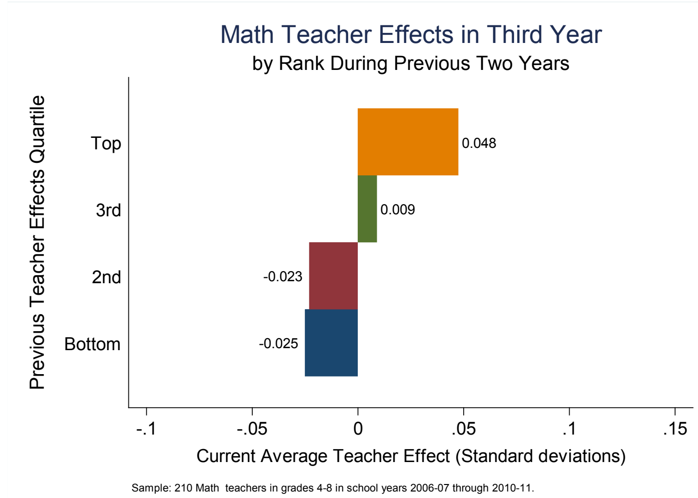
1. DISTRIBUTION OF TEACHERS BY VALUE-ADDED TEACHER EFFECT ESTIMATES

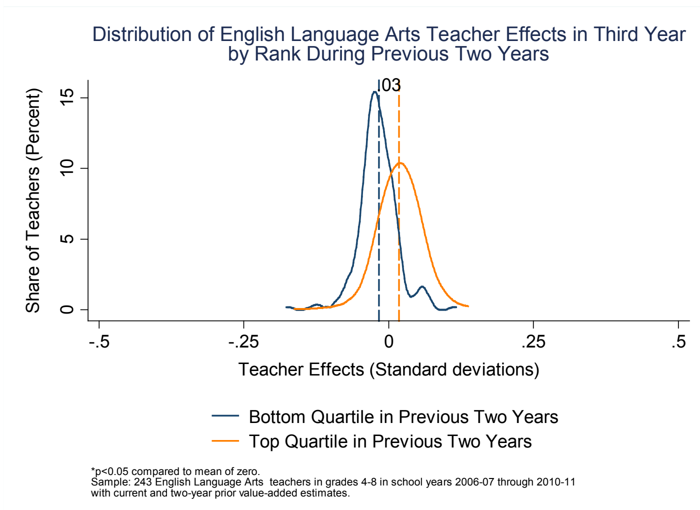
// Step 6: Create kdensity graph.

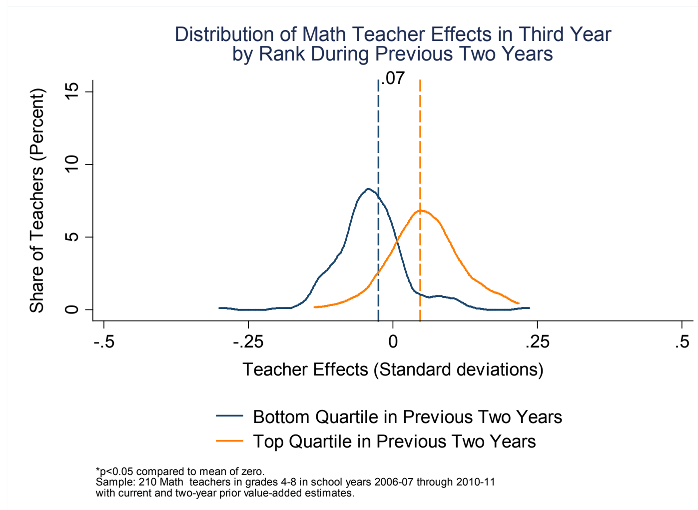
```
	#delimit ;

		histogram tre_`subj', kdensity percent width(.025) fcolor(none) lcolor(white)
		xline(`p10_`subj'', lpattern(dash) lcolor(green))
		xline(`p25_`subj'', lpattern(dash) lcolor(cranberry))
		xline(`p75_`subj'', lpattern(dash) lcolor(cranberry))
		xline(`p90_`subj'', lpattern(dash) lcolor(green))
		title("Distribution of `subj_title' Teacher Effects", margin(0 0 3 0))
		xtitle("`subj_title' Teacher Effects (Standard deviations)", margin(0 0 2
	2))
		ytitle("Share of Teachers (Percent)", margin(0 3 0 0))
		text(13 `p10_`subj'' "10th percentile", orientation(vertical))
		text(13 `p25_`subj'' "25th percentile", orientation(vertical))
		text(13 `p75_`subj'' "75th percentile", orientation(vertical))
		text(13 `p90_`subj'' "90th percentile", orientation(vertical))
		text(4  0 "``subj'diff1090'", color(green))
		text(7  0 "``subj'diff2575'", color(cranberry))
		legend(off)
		xscale(range(-.5(.25).5))
		xlabel(-.5(.25).5)
		yscale(range(0(5)15))
		ylabel(0(5)15, nogrid)
		${graph_pref}
		note("Sample: `samplesize' `subj_title'  teachers in grades $sample_grade in
	school years $sample_yr.", size(vsmall));

	#delimit cr

	graph save "${graphs}\D1_Evaluation_Kden_Overall_`caps_subj'.gph" , replace
	graph export "${graphs}\D1_Evaluation_Kden_Overall_`caps_subj'.emf" , replace

} // end of subject loop
}
```

2. PREDICTIVE POWER OF VALUE-ADDED IN FUTURE YEARS BASED ON PRIOR EFFECTIVENESS ESTIMATES









Purpose:

To show the extent to which prior value-added estimates of a teacher's effectiveness predict effectiveness in future years.

Required analysis file variables:

```
tid
school_year
curr2year_tre_math
curr2year_tre_ela
current_tre_math
current_tre_ela
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers included in value-added estimates.

Ask yourself:

- Value-added estimates are more likely to be accurate for teachers at the high and low ends of the distribution. What kinds of decisions would knowing which teachers consistently perform at the top and bottom ends of the distribution of value-added estimates help your agency make? How might this information be used for probationary teachers? For veteran teachers?

Potential further analyses:

- Restrict the sample to teachers in their third year at the agency to understand how performance of probationary teachers in their first two years compares to performance in the third year.

2. PREDICTIVE POWER OF VALUE-ADDED IN FUTURE YEARS BASED ON PRIOR EFFECTIVENESS ESTIMATES

Analytic technique: Group teachers into quartiles according to their value-added estimates in the prior two years of teaching, and then calculate the average value-added score of teachers in each of these quartiles in the current year.

```
/**** D. Evaluation****/
/*** 2. Create the predictive teacher effects analysis files and graphs ***/

if ${predictive_teacher_effects} == 1 {
```

// Step 1: Load the `Teacher_Year_Analysis` data file containing value-added estimates.

```
use "${analysis}/Teacher_Year_Analysis.dta", clear
```

// Step 2: Set the time series structure.

```
tsset tid school_year
```

// Step 3: Identify the most recent year a teacher is present in the data and tag as `year3`.

```
egen max_school_year = max(school_year), by(tid)
gen year3 = max_school_year == school_year
drop max_school_year
tab year3, m
```

// Step 4: Use lead operators to tag Years 2 and 1.

```
gen year2 = 0
bys tid: replace year2  = 1 if F.year3 == 1
gen year1 = 0
bys tid: replace year1 = 1 if F.year2 == 1
```

// Step 5: Keep only a balanced panel that includes teachers who have observations for all three years.

```
bys tid: egen balanced = max(year1)
keep if balanced == 1
drop balanced
codebook tid
```

// Loop through the subject areas (Math and ELA) to create the overall teacher evaluation analysis files.

```
foreach subj in math ela{
```

2. SEAMLESS AND DELAYED COLLEGE ENROLLMENT RATES BY HIGH SCHOOL

// Step 6: Store long titles in a variable to be used in the graphs.

```
if "`subj'" == "math" {
	local subj_title "Math"
	local caps_subj "Math"
}
if "`subj'" == "ela" {
	local subj_title "English/Language Arts"
	local caps_subj "ELA"
}
```

//Preserve the data file so that it can be used again in the loop.

```
preserve
```

// Step 7: Assign teachers to quartiles based on Year 1 and 2 pooled teacher effects.

```
xtile quart_yr1yr2_temp = curr2year_tre_`subj' if year2 == 1, nq(4)
bys tid: egen quart_yr1yr2 = max(quart_yr1yr2_temp)
label define quart 1 "Bottom" 2 "2nd" 3 "3rd" 4 "Top", replace
label values quart_yr1yr2 quart
label var quart_yr1yr2 "Quartiles of Teacher Effects in Previous Two Years"
```

// Step 8: Get quartile means in Year 3.

```
foreach n of numlist 1/4 {
	sum current_tre_`subj' if quart_yr1yr2 == `n' & year3 == 1,
	local quartmean`subj'`n' = r(mean)
}
```

// Step 9: Get the range between quartile 1 and quartile 4, and store it in a local variable.

```
local quartmeandiff`subj' = round(`quartmean`subj'4'-`quartmean`subj'1', .01)
```

// Step 10: Get the sample size.

```
count if year3 == 1 & !missing(current_tre_`subj')
local samplesize = r(N)
```

// Step 11: Check the significance of the model.

```
forvalues num = 1/4{
	gen quart_`num' = 0
	replace quart_`num' = 1 if quart_yr1yr2 == `num'
}
reg current_tre_`subj' quart_1 quart_2 quart_3 quart_4 if year3 == 1, nocons r
```

2. PREDICTIVE POWER OF VALUE-ADDED IN FUTURE YEARS BASED ON PRIOR EFFECTIVENESS ESTIMATES

// Step 12: Create a horizontal bar graph.

```
            #delimit ;

                  graph hbar current_tre_`subj' if year3 == 1,
                  over(quart_yr1yr2, descending)
                  asyvars showyvars blabel(bar, format(%9.3f) gap(*.5))
                  legend(off)
                  l1title("Previous Teacher Effects Quartile")
                  title("`subj_title' Teacher Effects in Third Year")
                  subtitle("by Rank During Previous Two Years")
                  ytitle("Current Average Teacher Effect (Standard deviations)", margin(0 0 2 2))
                  yscale(range(-.09 .15))
                  ylabel(-.10 (.05) .15, nogrid)
                  ${graph_pref}
                  note("Sample: `samplesize' `subj_title'  teachers in grades $sample_grade in
school            years $sample_yr.", size(vsmall));

            #delimit cr

            graph save "${graphs}\D2_Evaluation_Hbar_Alltchr_Predictive_`caps_subj'.gph" , replace

            graph export "${graphs}\D2_Evaluation_Hbar_Alltchr_Predictive_`caps_subj'.emf" ,
replace
```

// Step 13: Create an overlapping kdensity graph.

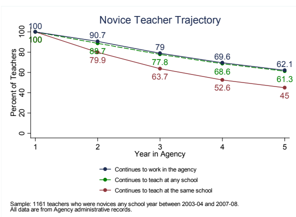
```
            #delimit ;

                  twoway (kdensity current_tre_`subj'
                  if quart_yr1yr2 == 1,
                  lcolor(navy)
                  area(`scale')
                  xline(`quartmean`subj'1', lpattern(dash) lcolor(navy)))
                  (kdensity current_tre_`subj'
                  if quart_yr1yr2 == 4,
                  lcolor(orange)
                  bwidth(.025)
                  xline(`quartmean`subj'4', lpattern(dash) lcolor(orange))),
                  yscale(range(0(5)15))
                  ylabel(0(5)15, nogrid)
                  xscale(range(-.5(.25).5))
                  xlabel(-.5(.25).5)
                  text(16 0 "`quartmeandiff`subj''")
                  title("Distribution of `subj_title' Teacher Effects in Third Year"
                  "by Rank During Previous Two Years", size(*.8) margin(0 0 3 0))
                  legend(order(1 2) rows(2) label(1 "Bottom Quartile in Previous Two Years")
                  label(2 "Top Quartile in Previous Two Years"))
                  xtitle("Teacher Effects (Standard deviations)", margin(0 0 2 2))
                  ytitle("Share of Teachers (Percent)", margin(0 3 0 0))
                  legend(symxsize(5) ring(1) region(lstyle(none) lcolor(none) color(none)))
                  ${graph_pref}
                  note("*p<0.05 compared to mean of zero."
                  "Sample: `samplesize' `subj_title'  teachers in grades $sample_grade in school
years $sample_yr"
```

2. PREDICTIVE POWER OF VALUE-ADDED IN FUTURE YEARS BASED ON PRIOR EFFECTIVENESS ESTIMATES

```
                    "with current and two-year prior value-added estimates.", size(vsmall))  ;

             #delimit cr

             graph save "${graphs}\D2_Evaluation_Kden_Alltchr_Predictive_`caps_subj'.gph", replace
             graph export "${graphs}\D2_evaluation_Kden_Alltchr_Predictive_`caps_subj'.emf",
replace

       restore
} // end subject loop
}
```

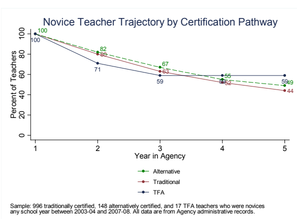
E. RETENTION

Schools invest a great deal of resources recruiting, developing, and retaining high-quality teachers. The analyses in the final step of the SDP Human Capital Diagnostic—Retention—reveal patterns of teachers transitioning to other schools within the system, moving to nonteaching positions, and exiting teaching in the agency altogether. The analyses reveal how retention patterns vary across school characteristics and among teachers with different performance ratings—value-added estimates, in this case. Education agencies typically offer various types of induction, mentoring, and support for first-year teachers. Nonetheless, novice teachers typically experience higher rates of turnover. For these reasons, some of the analyses in this section focus on the retention patterns of novice teachers.



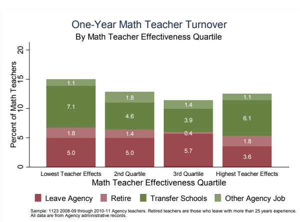
1. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS

Examines the basic novice teacher retention patterns in the agency.



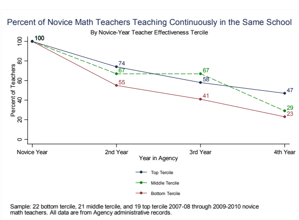
2. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS BY CERTIFICATION PATHWAY

Examines basic novice teacher retention patterns by certification pathway.



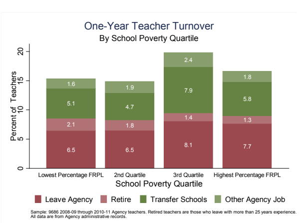
3. RETENTION BY TEACHER EFFECT QUARTILE

Determines whether the agency is retaining its highest-performing teachers.



4. FOUR-YEAR TRAJECTORY FOR NOVICES BY TEACHER EFFECTIVENESS TERCILE

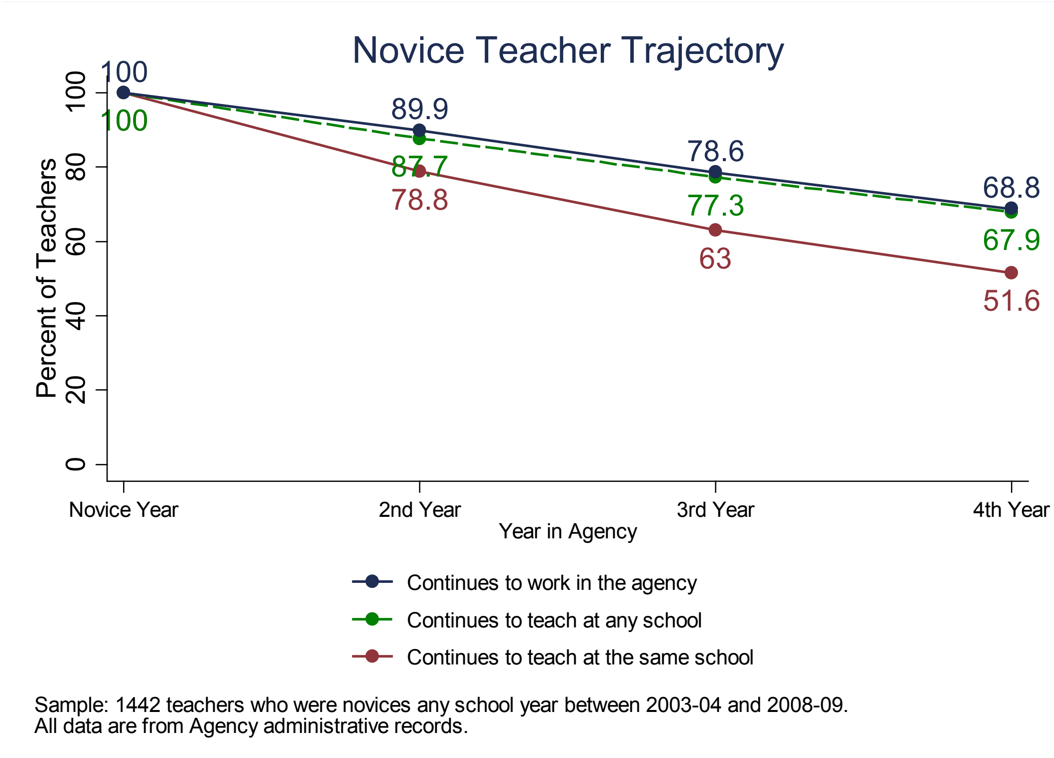
Examines whether the agency is retaining higher-performing early career teachers at higher rates than lower-performing early career teachers.



5. RETENTION BY SCHOOL POVERTY QUARTILE

Determines whether teacher retention patterns vary across school types.

1. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS



Purpose:
Examine basic novice teacher retention patterns for years in the agency.

Required analysis file variables:

```
tid
school_year
t_novice
t_transfer
t_stay
t_other_agency_job
t_leave
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers who begin their careers in the agency as novices, and for whom you can track career trajectory over the subsequent three school years.

Ask yourself:

- Does a sharp drop in retention occur in any year? If so, what might be driving turnover at this stage of a teacher's career?
- What types of support does the agency provide to novice teachers, and for how many years do early career teachers receive additional support?

Potential further analyses:

- Repeat the analysis, dividing novice teachers into those who start at high-poverty/high-need schools and those who start at lower-poverty schools.
- Repeat the analysis for teachers in the two subject areas that experience the most turnover or for teachers in high-need subject areas.
- Divide the sample into cohorts, and run the analysis to see if patterns hold over time.

1. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS

Analytic technique: Calculate the proportion of novice teachers who remain teaching at the same school, remain teaching in any school, and remain employed in the agency in the subsequent four school years.

```
/**** E. Retention ****/
/**** 1. Four Year Trajectory to Novice Teachers ****/

if $four_year_trajectory==1 {
```

// Step 1: Load the `Teacher_Year_Analysis_File` data file.

```
use "${analysis}\Teacher_Year_Analysis.dta", clear
```

// Step 2: Restrict the sample by including only teachers who began their careers in the agency as novices at least four school years prior to the most recent school year in the data.

```
keep if t_is_teacher==1
gen novice_before_2009 = (t_novice==1 & school_year<2010)
bys tid: egen max_novice_before_2010 = max(novice_before_2010)
keep if max_novice_before_2010==1
```

// Step 3: Create a variable that identifies the number of years each teacher was in the agency.

```
sort tid school_year
bys tid: gen year = _n
tab year
tab year if novice==1 // This should be 1 for everyone

label define year 1 "Novice" 2 "2nd Year" 3 "3rd Year" 4 "4th Year"
label values year year
```

// Step 4: Keep only the first three years.

```
keep if year <= 3
```

// Step 5: Identify the first-year teachers who leave their schools.

```
bys tid: egen temp_first_non_stay = min(year) if t_stay ==0
bys tid: egen first_non_stay = max(temp_first_non_stay)
replace t_stay = 0 if year > first_non_stay & !mi(first_non_stay)
```

// Step 6: Identify the first-year teachers who transferred schools.

```
bys tid: egen temp_first_transfer = min(year) if t_transfer==1
bys tid: egen first_transfer = max(temp_first_transfer)
replace t_transfer = 1 if year > first_transfer & !mi(first_transfer)
```

// Step 7: Create a variable that indicates whether the teacher is teaching anywhere in the agency.

```
gen teach_agency = (t_stay==1 | t_transfer==1)
gen agency_job = (t_stay==1 | t_transfer==1 | t_other_agency_job==1)
```

1. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS

// Step 8: Determine sample size of novice teachers.

```
sum tid if year==1
local sample = r(N)
```

// Step 9: Keep only the variables of interest for the analysis, and ensure that the data file is unique by teacher and school year.

```
keep tid year t_transfer t_stay t_leave t_other_agency_job teach_agency
agency_job
duplicates drop
```

// Step 10: Collapse for graphing.

```
collapse (sum) t_stay teach_agency agency_job, by(year)
```

// Step 11: Calculate the percentage of teachers in each category relative to the starting sample size.

```
gen total_count = `sample'
foreach var of varlist t_stay teach_agency agency_job {
	gen pct_`var' = round((`var'/total_count*100),.1)
}
```

// Step 12: Add 1 to the year variable to reflect the year in which the outcome variables take effect.

```
replace year = year + 1
local new_total_obs = _N + 1
set obs `new_total_obs'
replace year = 1 if mi(year)
```

// Step 13: Replace percentage of novice teachers as 100% in Year 1.

```
foreach var of varlist pct* {
	replace `var' = 100 if year == 1
}
```

// Step 14: Prepare the dataset for graphing.

```
keep year pct*
sort year
```

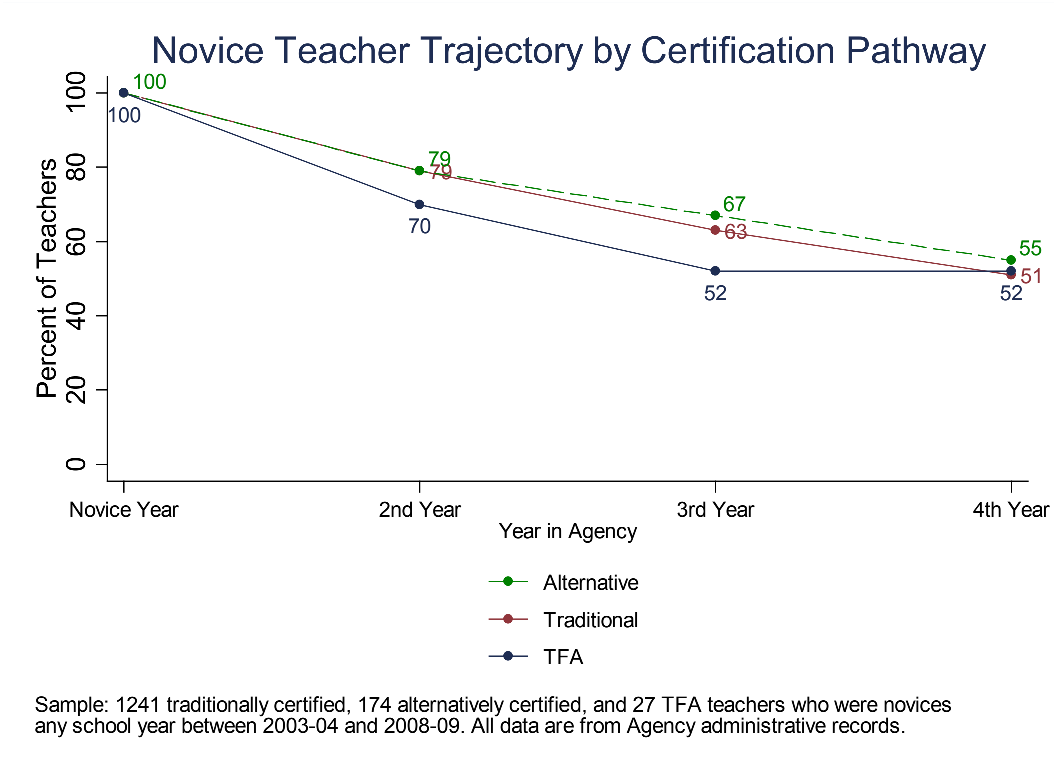
// Step 15: Create the graph.

```
#delimit ;
	twoway (connected pct_t_stay year, lcolor(maroon)
		lpattern(solid) msymbol(circle) mcolor(maroon) msize(medium)
		mlabel(pct_t_stay) mlabpos(6) mlabcolor(maroon) mlabsize(msmall))
		(connected pct_teach_agency year, lcolor(green)
		lpattern(dash) msymbol(circle) mcolor(green) msize(medium)
		mlabel(pct_teach_agency) mlabpos(6) mlabcolor(green) mlabsize(msmall))
		(connected pct_agency_job year, lcolor(dknavy)
		lpattern(solid) msymbol(circle) mcolor(dknavy) msize(medium)
```

1. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS

```
            mlabel(pct_agency_job) mlabpos(12) mlabcolor(dknavy) mlabsize(msmall)),
            title("Novice Teacher Trajectory")
            legend(col(1) order(3 2 1) position(6) size(small) symxsize(5) ring(1)
      region(lstyle(none)
            lcolor(none) color(none)) label(1 "Continues to teach at the same school")
            label(2 "Continues to teach at any school") label(3 "Continues to work in the
Agency"))
            ytitle("Percent of Teachers") xtitle("Year in Agency")
            yscale(range(0(20)100))
            ylabel(0(20)100, nogrid)
            xscale(range(1(1)3))
            xlabel(1 "Novice Year" 2 "2nd Year" 3 "3rd Year" 4 "4th Year", labsize(small))
            graphregion(color(white) fcolor(white) lcolor(white))
      plotregion(color(white) fcolor(white) lcolor(white))
            note("Sample: `sample' teachers who were novices any school year
      between 2003-04 and 2008-09."
            "All data are from ${agency_name} administrative records.", span pos(7)
size(small)) ;
      #delimit cr

graph export "${graphs}/E1_Four_Year_Novice_Trajectory.emf", replace
graph save "${graphs}/E1_Four_Year_Novice_Trajectory.gph", replace
}
```

2. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS BY CERTIFICATION PATHWAY



Purpose:
Examine basic novice teacher retention patterns by certification pathway.

Required analysis file variables:

```
tid
school_year
t_novice
t_stay
certification_pathway
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers who begin their careers in the agency as novices, and for whom you can track career trajectory over the subsequent three school years.

Ask yourself:

- How do retention rates vary over time for alternatively and traditionally certified teachers?
- How might this information guide recruitment decisions?

Potential further analyses:

- Restrict the sample to teachers who start at high-poverty/high-need schools.
- Examine the percentage of teachers who remain teaching in the same school over time by certification pathway.

2. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS BY CERTIFICATION PATHWAY

Analytic technique: Calculate the proportion of novice teachers in each certification pathway who remain teaching in the agency in the subsequent four school years.

```
/**** E. Retention ****/
/**** 2. Four-Year Trajectory of Novice Teachers by Certification Pathway****/

$four_year_trajectory_cert==1 {
```

// Step 1: Load the `Teacher_Year_Analysis_File` data file.

```
use "${analysis}\Teacher_Year_Analysis.dta", clear
```

// Step 2: Restrict the sample by including only teachers who began their careers in the agency as novices at least three school years prior to the most recent school year in the data.

```
keep if t_is_teacher==1
gen novice_before_2010 = (t_novice==1 & school_year<2010)
bys tid: egen max_novice_before_2010 = max(novice_before_2010)
keep if max_novice_before_2010==1
```

// Step 3: Create a variable that identifies the number of years each teacher was in the agency.

```
sort tid school_year
bys tid: gen year = _n
tab year
tab year if novice==1 // This should be 1 for everyone

label define year 1 "Novice" 2 "2nd Year" 3 "3rd Year" 4 "4th Year" 5 "5th Year"
label values year year
```

// Step 4: Keep only the first three years.

```
keep if year <= 3
```

// Step 5: Identify the first-year teachers who leave their schools.

```
bys tid: egen temp_first_non_stay = min(year) if t_stay ==0
bys tid: egen first_non_stay = max(temp_first_non_stay)
replace t_stay = 0 if year > first_non_stay & !mi(first_non_stay)
```

// Step 6: Identify the first-year teachers who transferred schools.

```
bys tid: egen temp_first_transfer = min(year) if t_transfer==1
bys tid: egen first_transfer = max(temp_first_transfer)
replace t_transfer = 1 if year > first_transfer & !mi(first_transfer)
```

// Step 7: Create a variable that indicates the teacher is teaching anywhere in the agency.

```
gen teach_agency = (t_stay==1 | t_transfer==1)
gen agency_job = (t_stay==1 | t_transfer==1 | t_other_agency_job==1)
```

2. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS BY CERTIFICATION PATHWAY

// Step 8: Determine sample size of novice teachers.

```
gen sample = .
forvalues x = 1/3 {
	unique tid if year==1 & t_certification_pathway==`x'
	replace sample = r(sum) if t_certification_pathway==`x'
	local sample_`x' = r(sum)
}
```

// Step 9: Keep only the variables of interest for the analysis, and ensure that the data file is unique by teacher and school year.

```
keep tid year t_transfer t_stay  agency_job t_certification_pathway sample
duplicates drop
```

// Step 10: Collapse for graphing.

```
collapse (sum) t_stay  (mean) sample, by(year t_certification_pathway)
```

// Step 11: Calculate the percentage of teachers in each category relative to the starting sample size.

```
foreach var of varlist t_stay  {
	gen pct_`var' = round(`var'/sample*100)
}
```

// Step 12: Add 1 to the year variable to reflect the year in which the outcome variables take effect.

```
replace year = year + 1

forvalues x = 1/3 {

	local new_total_obs = _N + 1
	set obs `new_total_obs'
	replace year = 1 if mi(year)
	replace t_certification_pathway = `x' if t_certification_pathway==.
}
```

// Step 13: Replace percentage of novice teachers as 100% in Year 1.

```
foreach var of varlist pct* {
	replace `var' = 100 if year == 1
}
```

// Step 14: Prepare the dataset for graphing.

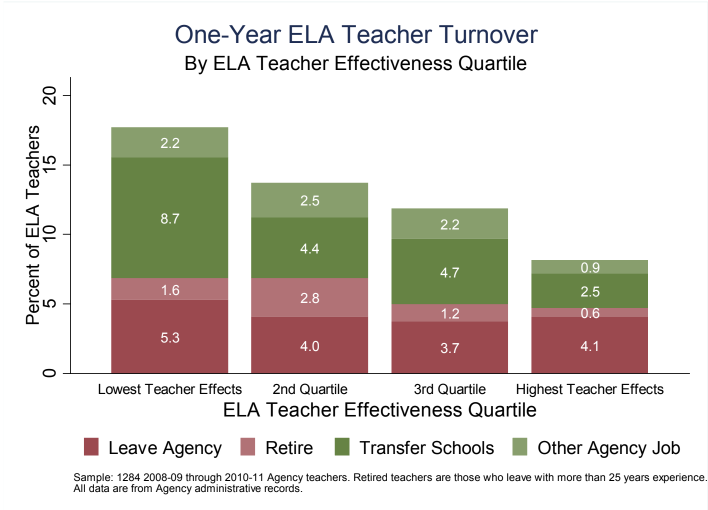
```
keep year pct* t_certification_pathway
sort year
```

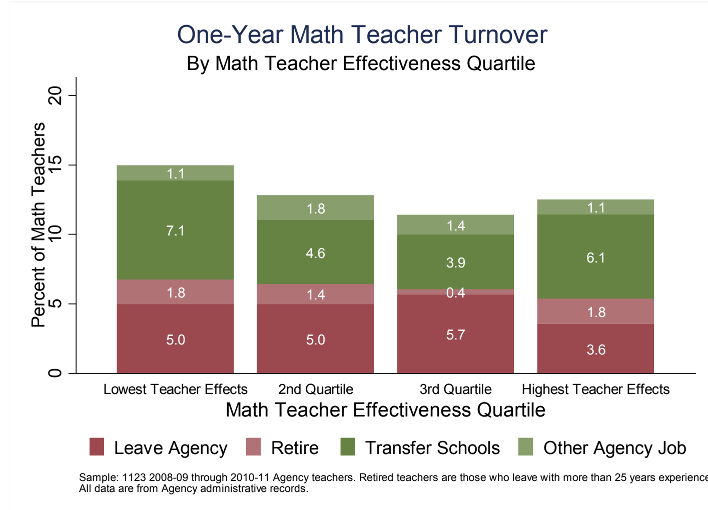
2. FOUR-YEAR TRAJECTORY OF NOVICE TEACHERS BY CERTIFICATION PATHWAY

// Step 15: Create the graph.

```
	#delimit ;
	twoway (connected pct_t_stay year if t_certification_pathway==1, lcolor(maroon)
		lpattern(solid) lwidth(vthin) msymbol(circle) mcolor(maroon) msize(small)
		mlabel(pct_t_stay) mlabpos(3) mlabcolor(maroon) mlabsize(small))
		(connected pct_t_stay year if t_certification_pathway==2, lcolor(green)
		lpattern(dash) lwidth(vthin) msymbol(circle) mcolor(green) msize(small)
		mlabel(pct_t_stay) mlabpos(2) mlabcolor(green) mlabsize(small))
	(connected pct_t_stay year if t_certification_pathway==3, lcolor(dknavy)
		lpattern(solid) lwidth(vthin) msymbol(circle) mcolor(dknavy) msize(small)
		mlabel(pct_t_stay) mlabpos(6) mlabcolor(dknavy) mlabsize(small)),
	title("Novice Teacher Trajectory by Certification Pathway")
		legend(col(1) order(2 1 3) position(6) size(small) symxsize(5) ring(1)
	region(lstyle(none)
		lcolor(none) color(none)) label(1 "Traditional")
		label(2 "Alternative") label(3 "TFA"))
		ytitle("Percent of Teachers") xtitle("Year in Agency", size(small))
		yscale(range(0(20)100))
		ylabel(0(20)100, nogrid)
		xscale(range(1(1)3))
		xlabel(1 "Novice Year" 2 "2nd Year" 3 "3rd Year" 4 "4th Year", labsize(small))
		graphregion(color(white) fcolor(white) lcolor(white)) plotregion(color(white)
fcolor(white) lcolor(white))
		note("Sample: `sample_1' traditionally certified, `sample_2' alternatively
	certified, and `sample_3' TFA teachers who were novices"
		"any school year between 2003-04 and 2007-08. All data are from ${agency_
name} administrative records.", span pos(7) size(small));
	#delimit cr
	gr_edit plotregion1.plot1.EditCustomStyle , j(1)
	style(label(textstyle(color(white))))

	graph export "${graphs}/E2_Four_Year_Novice_Trajectory_Certification_Path.emf",
replace
	graph save "${graphs}/E2_Four_Year_Novice_Trajectory_Certification_Path.gph", replace
}
```

3. RETENTION BY TEACHER EFFECT QUARTILE





Purpose:

Determine whether the agency is retaining its highest-performing teachers.

Required analysis file variables:

```
tid
school_year
t_leave
t_transfer
t_stay
t_other_agency_job
t_retire
curr2year_tre_math
curr2year_tre_ela
certification_pathway
```

Analysis-specific sample restrictions:

- Restrict the sample to teachers who have value-added estimates.

Ask yourself:

- How much of the district's turnover can be accounted for by the lowest-performing teachers?
- What are the characteristics of schools that teachers with low (or high) performance ratings transfer to?
- What might these patterns look like for teachers in non-tested grades and subjects?

Potential further analyses:

- Restrict the sample to teachers with fewer than four years of experience.
- Restrict the sample to be only teachers in high-poverty and/or low-performing schools.

3. RETENTION BY TEACHER EFFECT QUARTILE

Analytic technique: Within each teacher effectiveness quartile, calculate the proportion of teachers who transfer to teach at another school in the agency, transfer into a nonteaching position in the agency, and leave the agency altogether.

```
/**** E. Retention ****/
/**** 3. Retention by Teacher Effect Quartile****/

if $retention_by_VAM==1 {

foreach subject in math ela {
```

// Step 1: Load the `Teacher_Year_Analaysis_File` data file.

```
	use "${analysis}\Teacher_Year_Analysis.dta", clear
```

// Step 2: Create graph labels for each subject.

```
	if "`subject'" == "math" {
		local caps_subj "Math"
		local note "math"
	}
	if "`subject'" == "ela" {
		local caps_subj "ELA"
		local note "ELA"
	}
```

// Step 3: Restrict the analysis sample to include only teachers for whom retention data are not missing and who have value-added scores.

```
	egen nonmissing = rownonmiss(t_stay t_transfer t_other_agency_job t_leave
	t_retire)
	tab nonmissing, m
	keep if nonmissing == 5
	drop nonmissing
	keep if !mi(curr2year_tre_`subject')
```

// Step 4: Determine sample size of novice teachers.

```
	unique tid if !mi(curr2year_tre_`subject')
	local sample = r(N)
```

// Step 5: Create quartiles of value-added.

// 1. Create quartiles of value-added for each school year.

```
	forvalues year = 2009/2011 {
		xtile q`year' = curr2year_tre_`subject' if school_year == `year', nq(4)
	}
```

// 2. Create a single variable to indicate the value-added quartile.

```
	egen te_quartile = rowtotal(q*)
```

3. RETENTION BY TEACHER EFFECT QUARTILE

// Step 6: Keep the variables needed for the graph.

```
	keep tid school_year t_stay t_transfer t_other_agency_job t_leave t_retire
te_quartile
```

// Step 7: Collapse data for graphing.

```
	collapse (mean) t_stay t_transfer t_other_agency_job t_leave t_retire, by(te_quartile)
```

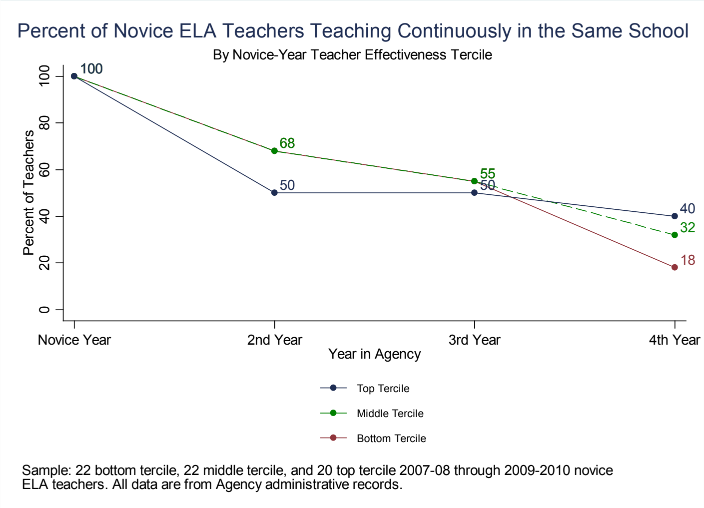
// Step 8: Multiply all values by 100.

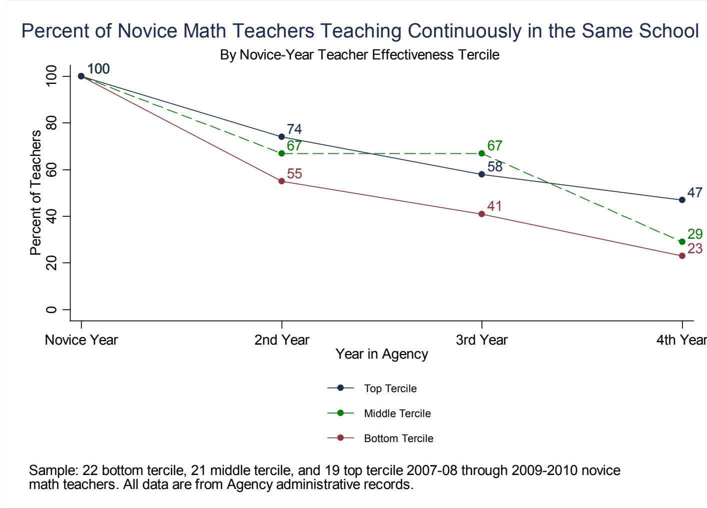
```
	foreach v in t_stay t_transfer t_other_agency_job t_leave t_retire{
		replace `v' = `v' * 100
	}
```

// Step 9: Create the graph.

```
	#delimit ;
		graph bar    t_leave t_retire t_transfer t_other_agency_job,
		over(te_quartile, relabel(1 "Lowest Teacher Effects" 2 "2nd Quartile" 3 "3rd
	Quartile" 4 "Highest Teacher Effects")
		label(labsize(small)) gap(20)) stack blabel(bar, size(small) gap(-2)
	position(center) format(%8.1f) color(white) )
		bar(1, color(maroon*0.9)) bar(2, color(maroon*0.7)) bar(3, color(forest_
green*0.9)) bar(4, color(forest_green*0.7))
		legend(position(6) order(1 2 3 4) cols(4) symxsize(3) ring(1)
	region(lstyle(none) lcolor(none) color(none))
		label(1 "Leave Agency") label(2 "Retire") label(3 "Transfer Schools")
	label(4 "Other Agency Job"))
		title("One-Year `caps_subj' Teacher Turnover", span)
		subtitle("By `caps_subj' Teacher Effectiveness Quartile", span)
		ytitle("Percent of `caps_subj' Teachers")
		b1title("`caps_subj' Teacher Effectiveness Quartile")
		yscale(range(0(5)20)) ylabel(0(5)20, nogrid)
		graphregion(color(white) fcolor(white) lcolor(white))
	plotregion(color(white) fcolor(white) lcolor(white))
		note("Sample: `sample' 2008-09 through 2010-11 ${agency_name} teachers.
Retired teachers are those who leave with more than 25 years experience."
		"All data are from ${agency_name} administrative records.", size(vsmall)) ;
	#delimit cr

	graph export "${graphs}\E3_Retention_by_VAM_`caps_subject'.emf", replace
graph save "${graphs}\E3_Retention_by_VAM_`caps_subject'.gph", replace
}
}
```

4. FOUR-YEAR TRAJECTORY FOR NOVICES BY TEACHER EFFECTIVENESS TERCILE





Purpose:

Examine whether the agency is retaining higher-performing early career teachers at higher rates than lower-performing early career teachers.

Required Analysis File Variables:

```
tid
school_year
t_novice
t_stay
current_tre_math
current_tre_ela
```

Analysis-Specific Sample Restrictions:

- Restrict the sample to teachers who begin their careers in the agency as novices, who have value-added estimates, and for whom you can track career trajectory over the subsequent four school years.

Ask Yourself:

- What systems and procedures are used to evaluate novice teachers? Are these standard across the agency? How are evaluation results used in teacher development and retention decisions? Do principals across the agency have access to evaluation data from other schools to inform hiring decisions?
- When high-performing teachers leave teaching in the agency, where do they go and why? What strategies could encourage retention? What strategies do districts with high retention rates of highly effective teachers use? Is the agency recruiting high-performing teachers out of the classroom to other jobs?
- Do principals feel empowered and know how to counsel out low-performing teachers? What information and training would principals need to act more intentionally on counseling out lower-performing teachers? Are principals held accountable for assembling and maintaining high-performing staff?

4. FOUR-YEAR TRAJECTORY FOR NOVICES BY TEACHER EFFECTIVENESS TERCILE

Analytic technique: Divide teachers into terciles based on their value-added estimate in the current year. Calculate the proportion of novice teachers in each tercile who remain teaching in any school in the agency in the subsequent four school years.

```
/**** E. Retention ****/
/**** 4. Four Year Trajectory for Novices by Teacher Effectiveness Tercile****/

if $four_year_trajectory_VAM==1 {

foreach subject in math ela {
```

// Step 1: Load the `Teacher_Year_Analaysis_File` data file.

```
		use "${analysis}\Teacher_Year_Analysis.dta", clear
```

// Step 2: Create graph labels for each subject.

```
	if "`subject'" == "math"{
		local caps_subj = "Math"
		local note = "math"
	}
	if "`subject'" == "ela"{
		local caps_subj = "ELA"
		local note = "ELA"
	}

	forvalues year = 2008/2010 {
		preserve
```

// Step 3: Identify cohort-specific novices.

```
			gen is_novice = 1 if school_year == `year' & t_novice == 1
			egen max_is_novice = max(is_novice), by(tid)
			keep if max_is_novice == 1
			drop max_is_novice is_novice
```

// Step 4: Classify novices into terciles by their first-year value-added estimates and keep only those teachers.

```
			xtile temp_tercile_`subject' = current_tre_`subject' if t_novice == 1, nq(3)
			egen tercile_`subject' = max(temp_tercile_`subject'), by(tid)
			keep if tercile_`subject' != .
			drop temp_tercile_`subject'
```

// Step 5: Restrict to only the first three years.

```
			keep if school_year >= `year' & school_year <= (`year' + 2)
```

// Step 6: Replace stay with zero for every year after the first time a teacher did not stay.

```
			egen temp_first_non_stay = min(school_year) if t_stay == 0, by(tid)
			egen first_non_stay = max(temp_first_non_stay), by(tid)
			replace t_stay = 0 if school_year > first_non_stay & first_non_stay != .
```

// Step 7: Keep relevant variables.

```
			keep tid school_year tercile_`subject' t_stay
```

4. FOUR-YEAR TRAJECTORY FOR NOVICES BY TEACHER EFFECTIVENESS TERCILE

// Step 8: Replace years with relative years.

```
                        replace school_year = school_year - `year' + 1
```

// Step 9: Save the tempfile.

```
                        tempfile tercile_retention_`year'
                        save `tercile_retention_`year''
                restore
        }
```

// Step 10: Combine the two cohorts.

```
        use `tercile_retention_2008', clear
        append using `tercile_retention_2009'
        append using `tercile_retention_2010'
```

// Step 11: Calculate sample size.

```
        gen sample = .
        forvalues x = 1/3 {
                unique tid if school_year == 1 & tercile_`subject' == `x'
                local sample_`x' = r(sum)
                replace sample = r(sum) if tercile_`subject' == `x'
        }
```

// Step 12: Collapse for graphing.

```
        collapse (sum) t_stay (mean) sample*, by(school_year tercile_`subject')
```

// Step 13: Generate outcome variables.

```
        gen pct_t_stay = round((t_stay/sample*100))
```

// Step 14: Shift years to reflect years in which outcome variables take effect; add reference group.

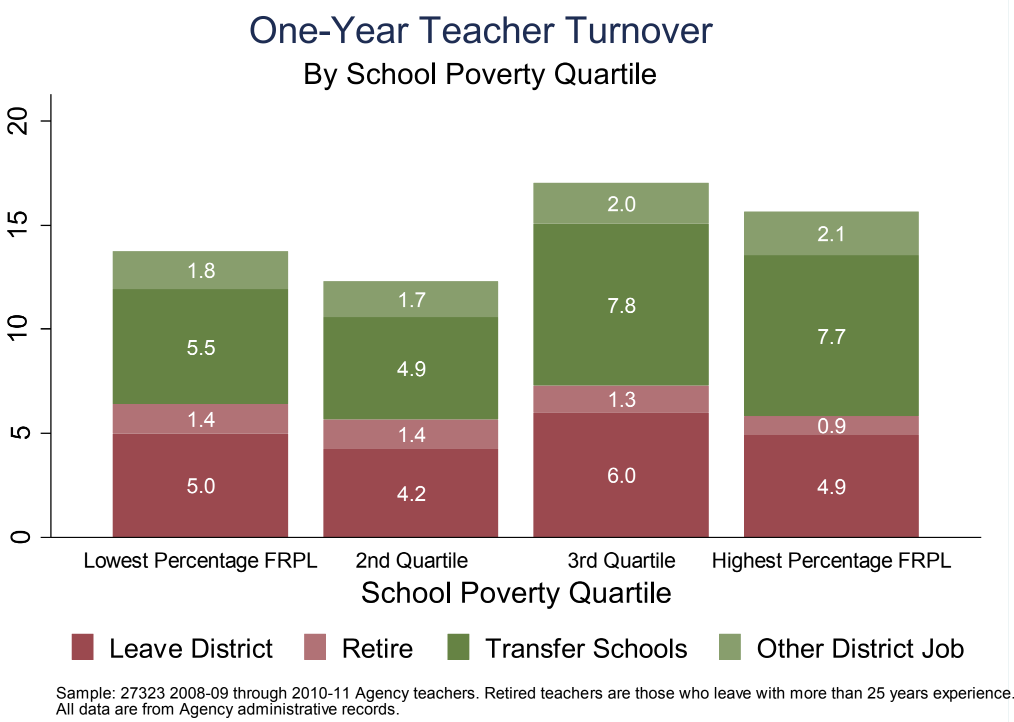
```
        replace school_year = school_year + 1
        local new_total_obs = _N + 3
        set obs `new_total_obs'
        replace school_year = 1 if school_year == .
        replace pct_t_stay = 100 if school_year == 1
        sort school_year tercile_`subject'
        forvalues x = 1/3{
                replace tercile_`subject' = `x' if _n == `x'
        }
        keep school_year tercile_`subject' pct
```

4. FOUR-YEAR TRAJECTORY FOR NOVICES BY TEACHER EFFECTIVENESS TERCILE

// Step 15: Graph and save.

```
	#delimit ;
		twoway (scatter  pct_t_stay school_year if tercile_`subject' == 1, connect(l) 
lcolor(maroon)
		lpattern(solid) lwidth(vthin) msymbol(circle) mcolor(maroon) msize(small)
		mlabel(pct_t_stay) mlabpos(2) mlabcolor(maroon) mlabsize(small))
		(scatter pct_t_stay school_year if tercile_`subject' == 2, connect(l) 
	lcolor(green)
		lpattern(dash)  lwidth(vthin) msymbol(circle) mcolor(green) mlabsize(small)
		msize(small) mlabel(pct_t_stay) mlabpos(2) mlabcolor(green))
		(scatter  pct_t_stay school_year if tercile_`subject' == 3, connect(l) 
	lcolor(dknavy)
		lpattern(solid)  lwidth(vthin) msymbol(circle) mcolor(dknavy) msize(small)
		mlabel(pct_t_stay) mlabpos(2) mlabcolor(dknavy) mlabsize(small)),
		title("Percent of Novice `caps_subj' Teachers Teaching Continuously in the 
	Same School", size(msmall) span)
		subtitle("By Novice-Year Teacher Effectiveness Tercile", size(small) span)
		legend(col(1) order(3 2 1) size(vsmall) symxsize(5) ring(1) 
	region(lstyle(none) lcolor(none) color(none))
		label(1 "Bottom Tercile")
		label(2 "Middle Tercile")
		label(3 "Top Tercile"))
		ytitle("Percent of Teachers", size(small)) xtitle("Year in Agency", 
	size(small))
		yscale(range(0(20)100))
		ylabel(0(20)100, nogrid labsize(small))
		xscale(range(1(1)4))
		xlabel(1 "Novice Year" 2 "2nd Year" 3 "3rd Year" 4 "4th Year", labsize(small))
		graphregion(color(white) fcolor(white) lcolor(white))
		plotregion(color(white) fcolor(white) lcolor(white))
		note("Sample: `sample_1' bottom tercile, `sample_2' middle tercile, and 
`sample_3' top tercile 2007-08 through 2009-2010 novice"
		"`note' teachers. All data are from ${agency_name} administrative records.", 
	size(small) span pos(7));

	#delimit cr

	graph export "${graphs}\E4_Retention_by_VAM_`subject'.emf", replace
	graph save "${graphs}\E4_Retention_by_VAM_`subject'.gph", replace

} // end subject loop
}
```

5. RETENTION BY SCHOOL POVERTY QUARTILE



Purpose:
Determine whether teacher retention patterns vary across school types.

Required analysis file variables:

```
tid
school_year
t_transfer
t_leave
t_other_agency_job
t_stay
t_retire
school_poverty_quartile
```

Analysis-specific sample restrictions:

- Restrict the sample to include only teachers for whom retention data are not missing.

Ask yourself:

- What percentages of teachers turnover in each quartile of teacher poverty?
- Do these teachers appear to transfer to other schools, leave the district, or move to other district jobs?

Potential further analyses:

- Consider examining which schools teachers are transferring to for different school poverty quartiles.
- Examine transfers from teacher to principal status and vice versa.

5. RETENTION BY SCHOOL POVERTY QUARTILE

Analytic technique: Within each school poverty quartile, calculate the proportion of teachers who remain teaching at the same school, transfer to teach at another school in the system, transfer into a nonteaching position in the system, and leave the system altogether.

```
/**** E. Retention ****/
/**** 5. Retention by School Poverty Quartile****/

if $retention_by_school_poverty==1 {

use "${analysis}\Teacher_Year_Analysis.dta", clear
merge m:1 school_code using "${clean}\School_Clean.dta", keep(3) nogen
```

// Step 1: Restrict the analysis sample to include only teachers for whom retention data are not missing.

```
	egen nonmissing = rownonmiss(t_stay t_transfer t_other_agency_job t_leave
	t_retire)
	tab nonmissing, m
	keep if nonmissing == 5
	drop nonmissing

	**To remove later, when school poverty is fixed**
	drop if mi(school_poverty_quartile)
```

// Step 2: Determine sample size of novice teachers.

```
	unique tid
	local sample = r(N)
```

// Step 3: Keep the variables needed for the graph.

```
	keep tid school_year t_stay t_transfer t_other_agency_job t_leave t_retire
school_poverty_quartile
```

// Step 4: Collapse data for graphing.

```
	collapse (mean) t_stay t_transfer t_other_agency_job t_leave t_retire,
by(school_		poverty_quartile)
```

// Step 5: Multiply all values by 100.

```
	foreach v in t_stay t_transfer t_other_agency_job t_leave t_retire{
		replace `v' = `v' * 100
	}
```

5. RETENTION BY SCHOOL POVERTY QUARTILE

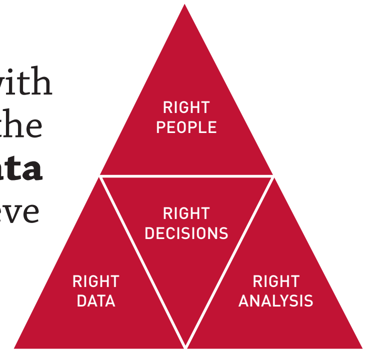
// Step 6: Create the graph.

```
	#delimit ;
		graph bar	t_leave t_retire t_transfer t_other_agency_job,
		over(school_poverty_quartile, relabel(1 "Lowest Percentage FRPL" 2 "2nd 
Quartile" 3 "3rd Quartile" 4 "Highest Percentage FRPL")
		label(labsize(small)) gap(20)) stack blabel(bar, size(small) gap(-2) 
	position(center) format(%8.1f) color(white) )
		bar(1, color(maroon*0.9)) bar(2, color(maroon*0.7)) bar(3, color(forest_
green*0.9)) bar(4, color(forest_green*0.7))
		legend(position(6) order(1 2 3 4) cols(4) symxsize(3) ring(1) 
	region(lstyle(none) lcolor(none) color(none))
		label(1 "Leave Agency") label(2 "Retire") label(3 "Transfer Schools") 
label(4 "Other Agency Job"))
		title("One-Year `caps_subj' Teacher Turnover", span)
		subtitle("By `caps_subj' Teacher Effectiveness Quartile", span)
		ytitle("Percent of `caps_subj' Teachers")
		b1title("`caps_subj' Teacher Effectiveness Quartile")
		yscale(range(0(5)20)) ylabel(0(5)20, nogrid)
		graphregion(color(white) fcolor(white) lcolor(white))
		plotregion(color(white) fcolor(white) lcolor(white))
		note("Sample: `sample' 2008-09 through 2010-11 ${agency_name} teachers. 
Retired teachers are those who leave with more than 25 years of experience."
		"All data are from ${agency_name} administrative records.", size(vsmall));
	#delimit cr

	graph export "${graphs}\E5_Retention_by_School_Poverty_Quartile.emf", replace
	graph save "${graphs}\E5_Retention_by_School_Poverty_Quartile.gph", replace
}
```

The Strategic Data Project

OVERVIEW

The Strategic Data Project (SDP), housed at the Center for Education Policy Research at Harvard University, partners with school districts, school networks, and state agencies across the United States. **Our mission is to transform the use of data in education to improve student achievement.** We believe that with the right people, the right data, and the right analyses, we can improve the quality of strategic policy and management decisions.



SDP AT A GLANCE

56 AGENCY PARTNERS
34 SCHOOL DISTRICTS
12 STATE EDUCATION DEPARTMENTS
3 CHARTER SCHOOL ORGANIZATIONS
7 NONPROFIT ORGANIZATIONS

107 FELLOWS
65 CURRENT
42 ALUMNI

CORE STRATEGIES

1. Building a network of top-notch data strategists who serve as fellows for two years with our partners

2. Conducting rigorous diagnostic analyses of teacher effectiveness and college-going success using existing agency data

3. Disseminating our tools, methods, and lessons learned to the education sector broadly

SDP DIAGNOSTICS

SDP's second core strategy, conducting rigorous diagnostic analyses using existing agency data, focuses on two core areas: (1) college-going success and attainment for students, and (2) human capital (primarily examining teacher effectiveness).

The diagnostics are a set of analyses that frame actionable questions for education leaders. By asking questions such as "How well do students transition to postsecondary education?" or "How successfully is an agency recruiting effective teachers?" we support education leaders to develop a deep understanding of student achievement in their agency.

ABOUT THE SDP TOOLKIT FOR EFFECTIVE DATA USE

SDP's third core strategy is to disseminate our tools, methods, and lessons learned to education agencies broadly. This toolkit is meant to help analysts in all education agencies collect data and produce meaningful analyses in the areas of college-going success and teacher effectiveness. Notably, the analyses in this release of our toolkit primarily support questions related to college-going success. The data collection (Identify) and best practices (Adopt) stages of the toolkit, however, are applicable to any sort of diagnostic and convey general data use guidelines valuable to any analysts interested in increasing the quality and rigor of their analyses.

Center for Education Policy Research
HARVARD UNIVERSITY

CENTER FOR EDUCATION POLICY RESEARCH
STRATEGIC DATA PROJECT
50 CHURCH ST., 4TH FLOOR, CAMBRIDGE, MA 02138
VOX 617.496.1563
FAX 617.495.2614
WWW.GSE.HARVARD.EDU/SDP