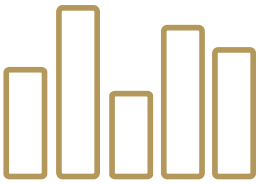


STRATEGIC DATA PROJECT

ANALYZE: COLLEGE-GOING SUCCESS ANALYSIS GUIDE

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 College-Going

Connect: Data Linking Guide for College-Going

Analyze: College-Going Success Analysis Guide

Adopt: Coding Style Guide

SDP Stata Glossary

VERSION: 1.2
Last Modified: December 22nd, 2014

| Authored by Todd Kawakita and the SDP Research Team

4. Analyze: College-Going Success Analysis Guide

Conduct analyses that help answers key questions in your agency.

Analyze: College-Going Success Analysis Guide is a set of step-by-step instructions to help you generate data visualizations about student pathways through high school and college. Through **Analyze**, your previous work identifying, cleaning, and connecting data will generate actual analyses to inform decision making in your agency!

COLLEGE-GOING SUCCESS ANALYSIS GUIDE

So far you have identified, cleaned, and connected your data into a single analysis file named CG_Analysis. **Analyze** uses this final analysis file to generate a number of analyses along the student pipeline through high school and college.

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 high school completion and college-going success 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;
- **Possible Next Steps or Action Plans:** further analyses you may conduct to understand underlying causes or interventions needed (this section is included in some but not all analyses)
- **Analytic Technique:** how to produce the analysis step-by-step using your analysis file and code in Stata.

Analysis-Specific Sample Restrictions

One of the most important decisions in running each analysis is defining the sample. Each analysis corresponds to a different part of the education pipeline and as a result require different cohorts of students.

If you are using the synthetic data we have provided (available for download at **www.gse.harvard.edu/sdp/tools**), the sample restrictions have been predefined and are included on the next page.

If you are using your own agency data, change these sample restrictions based on your data.

Note that you will have to run these sample restrictions at the beginning of your Analyze do file so they will feed into the rest of your Stata code.

```
/*** Sample Restrictions ***/
// Agency name
global agency_name "Agency"

// Ninth grade cohorts you can observe persisting to the second year of college
global chrt_ninth_begin_persist_yr2 = 2005
global chrt_ninth_end_persist_yr2 = 2005

// Ninth grade cohorts you can observe graduating high school on time
global chrt_ninth_begin_grad = 2005
global chrt_ninth_end_grad = 2006

// Ninth grade cohorts you can observe graduating high school one year late
global chrt_ninth_begin_grad_late = 2005
global chrt_ninth_end_grad_late = 2005

// High school graduation cohorts you can observe enrolling in college the fall after
graduation
global chrt_grad_begin = 2008
global chrt_grad_end = 2009

// High school graduation cohorts you can observe enrolling in college two years after hs
graduation
/*global chrt_grad_begin_delayed = 2008
global chrt_grad_end_delayed = 2008*/
```

Based on the sample data, you will have no more than two cohorts (sometimes only one) for analysis. If your own agency data is more extensive, you may decide to aggregate results for three or four cohorts to report your results. This decision depends on 1) how much historical data you have (you may only have two cohorts of data) and 2) what balance to strike between reliability and averaging away information on recent trends. We suggest you average results for the last three cohorts to take advantage of larger sample sizes and improve reliability. However, if you have data for more than three cohorts, you may decide to not average data out for fear of losing information about trends and recent changes in your agency.

Strategic Performance Indicators (SPIs)

Three of the analyses in the College Enrollment section of **Analyze** include the Strategic Performance Indicators (SPIs) SDP has released to provide deeper insight into the college-going performance of educational systems. These SPIs were produced using data from a number of SDP's partner agencies. You will be able to conduct these analyses yourself through **Analyze**. You can read more about the SPIs at **http://www.gse.harvard.edu/sdp/spi**.

Summary

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

- Used your final analysis file from **Connect** to generate many different analyses to display student outcomes along the education pipeline;
- Obtained new and confirmatory information about high school and college-going outcomes 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?

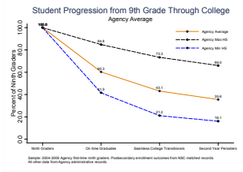
As a final note, the analyses presented here do not capture all of our research team's efforts to understand college-going success. We believe the analyses presented are the most widely applicable to drive discussions about change. Moreover, we believe these analyses serve as a model to seek answers about postsecondary success.

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**.

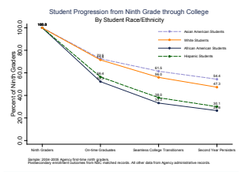
Map of Analyses

A. Attainment Along the Education Pipeline

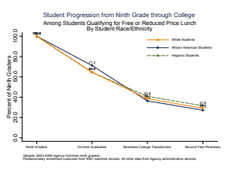
An overall picture of student attainment from high school entry through high school graduation, college enrollment, and college persistence:



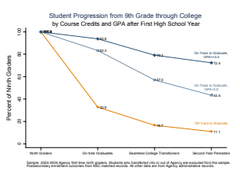
1. Overall Progression
(p. 9)



2. Progression by Student Race/Ethnicity
(p. 13)



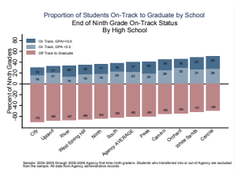
3. Progression by Student Race/Ethnicity, Among FRPL-Eligible Students
(p. 17)



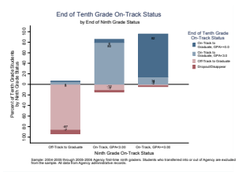
4. Progression by Students' On-Track Status After Ninth Grade
(p. 21)

B. Ninth to Tenth Grade Transition

An examination of the ninth to tenth grade transition in a student's high school career, with attention to credit accumulation for graduation. Research cites academic performance in ninth grade as a strong predictor of high school graduation:



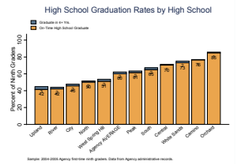
1. Proportion of Students On-Track at the End of Ninth Grade, By High School
(p. 26)



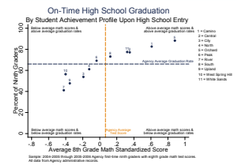
2. Ninth to Tenth Grade Transition by On-Track Status
(p. 31)

C. High School Graduation

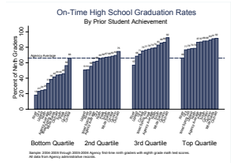
High school graduation is a crucial step towards higher education. Analyzing this outcome across schools and student subgroups illustrates how certain high schools may influence graduation:



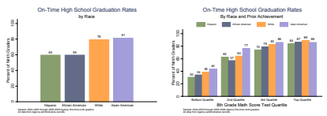
1. High School Completion Rates By School
(p. 35)



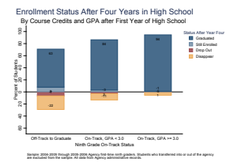
2. High School Completion Rates By Average 8th Grade Achievement
(p. 38)



3. High School Completion Rates By 8th Grade Achievement Quartiles
(p. 41)



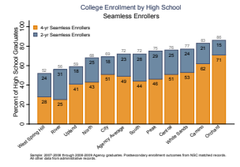
4. Racial Gaps In Completion Overall And By 8th Grade Achievement Quartiles
(p. 44)



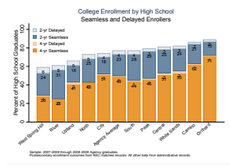
5. Enrollment Outcome In Year 4 By On-Track Status At The End Of Ninth Grade
(p. 48)

D. College Enrollment

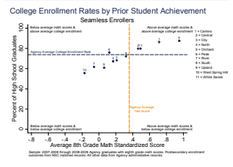
High schools are responsible for preparing students to enter and continue through college. In these analyses you explore college going rates across high schools:



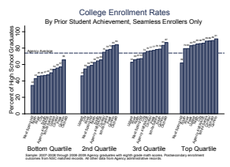
1. College Enrollment Rates by High School
(p. 54)



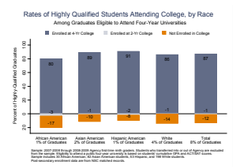
2. Enrollment Selectivity at Four-Year Colleges by High School
(p. 57)



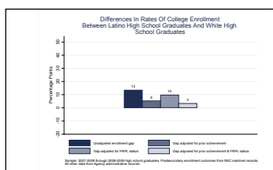
3. College Enrollment Rates by Average 8th Grade Achievement
(p. 60)



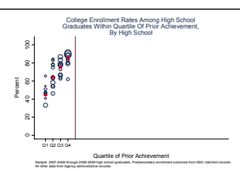
4. College Enrollment Rates by 8th Grade Achievement Quartiles
(p. 64)



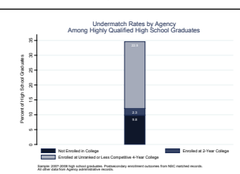
5. Rates of Non-Enrollment among Graduates Highly Qualified to Attend Four-Year Colleges
(p. 68)



6. Gaps in Rates of College Enrollment between Latino High School Graduates and White school Graduates
(p. 73)



7. College Enrollment Rates by 8th Grade Achievement Quartiles - Bubbles
(p. 77)

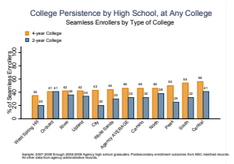


8. Undermatch Rates Among Highly Qualified High School Graduates
(p. 80)

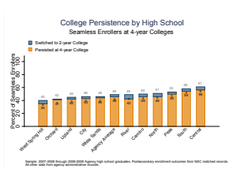
Strategic Performance Indicators

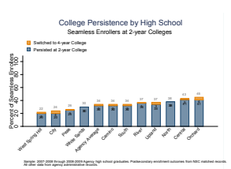
E. College Persistence

Enrollment is important but college persistence is also essential. In this section you explore patterns of persistence to the second year of college to identify indicators of student progress towards degree attainment:



1. Persistence Rates to the Second Year of College by High School
(p. 86)





2. Persistence Across Two-Year and Four-Year Colleges
(p. 90)

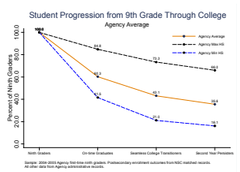
3. Top-Enrolling Colleges/ Universities of Agency Graduates
(p. 94)

A. Attainment along the Education Pipeline

Attainment along the Education Pipeline analyses summarize student attainment from ninth grade through college using three milestones: 1) on-time high school completion, 2) seamless college transition, and 3) persistence to the second year of college.

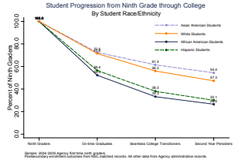
Through these analyses, you identify drop-offs along the education pipeline for students as a group and as subgroups. For different subgroups, these analyses illuminate disparities in college attainment by race, family income, high school attended, and academic achievement. A steep decline in college enrollment from high school completion date for specific subgroups may indicate barriers to college access. On the other hand, a steep decline from initial college enrollment to second-year persistence might suggest students were not prepared for rigorous college coursework during high school.

The analyses exploring attainment along the pipeline include:



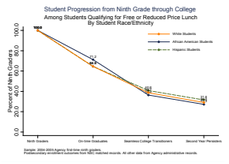
1. OVERALL PROGRESSION

Tracks the overall percent of ninth graders who complete high school on-time, seamlessly enroll in college, and persist to the second year of college.



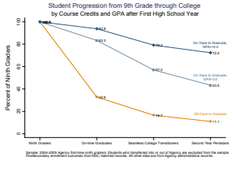
2. PROGRESSION BY STUDENT RACE/ETHNICITY

Tracks the percent of ninth graders of different races/ethnicities who complete high school on-time, seamlessly enroll in college, and persist to the second year of college.



3. PROGRESSION BY STUDENT RACE/ETHNICITY, AMONG FRPL-ELIGIBLE STUDENTS

Tracks the percent of ninth graders of different races/ethnicities who ever qualified for free or reduced price lunch,complete high school on-time, seamlessly enroll in college, and persist to the second year of college.

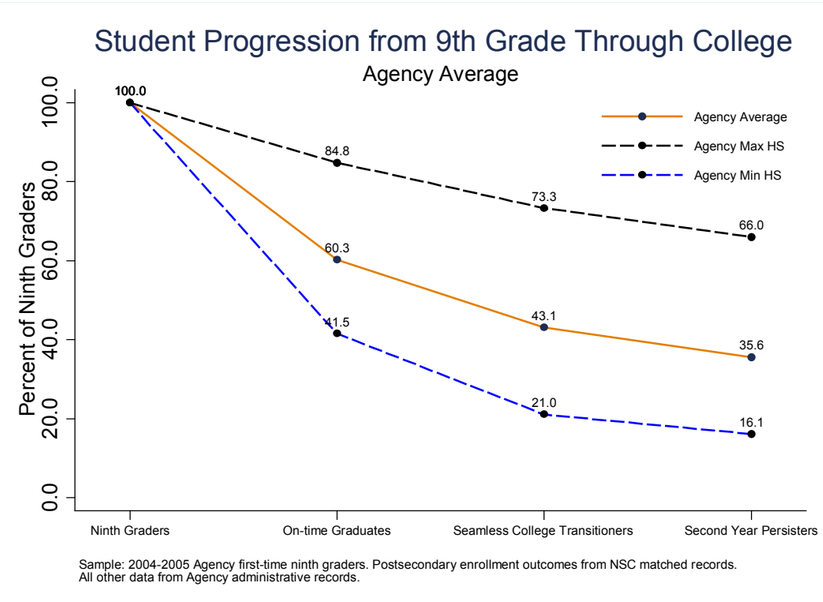


4. PROGRESSION BY STUDENTS' ON-TRACK STATUS AFTER NINTH GRADE

Tracks the percent of ninth graders at different levels of being on-track for graduation who complete high school on-time, seamlessly enroll in college, and persist to the second year of college.

A. Attainment along the Education Pipeline

1. OVERALL PROGRESSION



Purpose: This analysis tracks the overall percent of ninth graders who complete high school on-time, seamlessly enroll in college, and persist to the second year of college. To examine the range of attainment at each milestone, the minimum and maximum values of any high school are shown.

Required Analysis File Variables:

sid
chrt_ninth
first_hs_name
ontime_grad
enrl_1oct_ninth_yr1_any
enrl_1oct_ninth_yr2_any

Analysis-Specific Sample Restrictions: Keep students in ninth grade cohorts for which persistence to the second year of college can be reported.

Ask Yourself

- Do you notice drop-offs along the pipeline?
- Are differences in agency maxima and minima at different points along the pipeline surprising? What might be different about these high schools?
- Are your numbers in line with agency-reported figures in other publicly available reports? What might account for differences?

1. OVERALL PROGRESSION

Analytic Technique: Calculate the proportion of first-time ninth graders that progress to each step along the education pipeline.

```
/**** A. Attainment along the Education Pipeline ****/
/**** 1. Overall Progression ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe persisting to the second year of college

```
local chrt_ninth_begin = ${chrt_ninth_begin_persist_yr2}
local chrt_ninth_end = ${chrt_ninth_end_persist_yr2}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end')
```

// Step 3: Create variables for the outcomes "regular diploma recipients", "seamless transitioners" and "second year persisters"

```
gen grad = (!mi(chrt_grad) & ontime_grad == 1)
gen seamless_transitioners_any = (enrl_1oct_ninth_yr1_any == 1 & ontime_grad == 1)
gen second_year_persisters = (enrl_1oct_ninth_yr1_any == 1 & enrl_1oct_ninth_yr2_any == 1
& ontime_grad == 1)
```

// Step 4: Create agency-level average outcomes

```
// 1. Preserve the data (to work with the data in its existing structure later on)
preserve
// 2. Calculate the mean of each outcome variable by agency
    collapse (mean) grad seamless_transitioners_any second_year_persisters (count) N = sid
// 3. Create a string variable called school_name equal to "${agency_name} Average"
    gen school_name = "${agency_name} AVERAGE"
// 4. Save this data as a temporary file
    tempfile agency_level
    save `agency_level'
// 5. Restore the data to the original form
restore
```

// Step 5: Create school-level maximum and minimum outcomes

```
// 1. Create a variable school_name that takes on the value of students' first high school attended
gen school_name = first_hs_name
// 2. Calculate the mean of each outcome variable by first high school attended
collapse (mean) grad seamless_transitioners second_year_persisters (count) N = sid,
by(school_name)
// 3. Identify the agency maximum values for each of the three outcome variables
preserve
    collapse (max) grad seamless_transitioners_any second_year_persisters (count) N
    gen school_name = "${agency_name} MAX HS"
    tempfile agency_max
    save `agency_max'
restore
```


1. OVERALL PROGRESSION

```
// 4. Identify the agency minimum values for each of the three outcome variables
preserve
    collapse (min) grad seamless_transitioners_any second_year_persisters (count) N
    gen school_name = "${agency_name} MIN HS"
    tempfile agency_min
    save `agency_min'
restore

// 5. Append the three tempfiles to the school-level file loaded into Stata
append using `agency_level'
append using `agency_max'
append using `agency_min'
```

// Step 6: Format the outcome variables so they read as percentages in the graph

```
foreach var of varlist grad seamless_transitioners_any second_year_persisters {
    replace `var' = (`var' * 100)
    format `var' %9.1f
}
```

// Step 7: Reformat the data file so that one variable contains all the outcomes of interest

```
// 1. Create 4 observations for each school: ninth grade, hs graduation, seamless college transition and second-year
persistence
foreach i of numlist 1/4 {
    gen time`i' = `i'
}
// 2. Reshape the data file from wide to long
reshape long time , i(school_name N)
drop _j
// 3. Create a single variable that takes on all the outcomes of interest
bysort school_name: gen outcome = 100 if time == 1
bysort school_name: replace outcome = grad if time == 2
bysort school_name: replace outcome = seamless_transitioners_any if time == 3
bysort school_name: replace outcome = second_year_persisters if time == 4
format outcome %9.1f
```

// Step 8: Prepare to graph the results

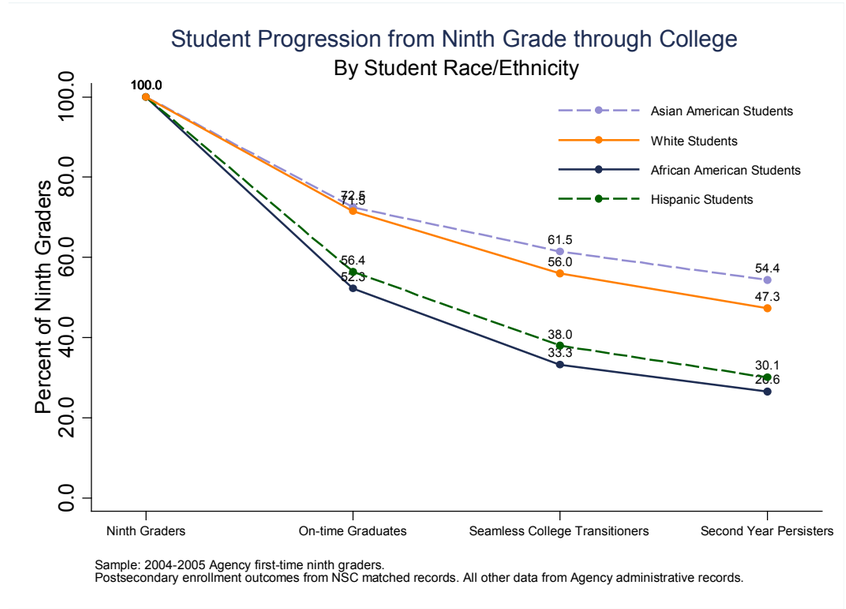
```
// 1. Label the outcome
label define outcome 1 "Ninth Graders" 2 "On-time Graduates" ///
3 "Seamless College Transitioners" 4 "Second Year Persisters"
label values time outcome
// 2. Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

1. OVERALL PROGRESSION

// Step 9: Graph the results

```
#delimit ;
twoway (connected outcome time if school_name == "${agency_name} AVERAGE",
    sort lcolor(dkorange) mlabel(outcome) mlabc(black) mlabs(vsmall) mlabp(12)
    mcolor(dknavy) msymbol(circle) msize(small))
    (connected outcome time if school_name == "${agency_name} MAX HS", sort lcolor(black)
    lpattern(dash) mlabel(outcome) mlabs(vsmall) mlabp(12) mlabc(black)
    mcolor(black) msize(small))
    (connected outcome time if school_name == "${agency_name} MIN HS", sort lcolor(blue)
    lpattern(dash) mlabel(outcome) mlabs(vsmall) mlabp(12) mlabc(black)
    mcolor(black) msize(small)),
title("Student Progression from 9th Grade Through College")
    subtitle("${agency_name} Average", size(medsmall))
    xscale(range(.8(.2)4.2))
    xtitle("") xlabel(1 2 3 4 , valuelabels labsize(vsmall))
    ytitle("Percent of Ninth Graders")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
legend(col(1) position(2) size(vsmall)
    label(1 "${agency_name} Average")
    label(2 "${agency_name} Max HS")
    label(3 "${agency_name} Min HS")
    ring(0) region(lpattern(none) lcolor(none) fcolor(none)))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders. Postsecondary
enrollment outcomes from NSC matched records." "All other data from ${agency_name}
administrative records.", size(vsmall));
#delimit cr
graph export "A1_Overall_Progression.emf", replace
graph save "A1_Overall_Progression.gph", replace
}
```


2. PROGRESSION BY STUDENT RACE/ETHNICITY



Purpose: This analysis tracks the percent of ninth graders of different races/ethnicities who complete high school on-time, seamlessly enroll in college, and persist to the second year of college.

Required Analysis File Variables:

- sid
- race_ethnicity
- chrt_ninth
- ontime_grad
- enrl_1oct_ninth_yr1_any
- enrl_1oct_ninth_yr2_any

Analysis-Specific Sample Restrictions:

- Keep students in ninth grade cohorts for which persistence to the second year of college can be reported.
- Restrict the sample to include students from the most representative racial/ethnic sub-groups.

Ask Yourself

- Which races/ethnicities face larger drop-offs along the pipeline?
- Might certain groups face different barriers to progressing along the education pipeline?

2. PROGRESSION BY STUDENT RACE/ETHNICITY

Analytic Technique: Calculate the proportion of first-time ninth graders that progress to each // Step along the education pipeline.

/**** A. Attainment along the Education Pipeline ****/
/**** 2. Progression by Student Race/Ethnicity ****/
{

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe persisting to the second year of college

```
local chrt_ninth_begin = ${chrt_ninth_begin_persist_yr2}
local chrt_ninth_end = ${chrt_ninth_end_persist_yr2}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end')
```

// Step 3: Create variables for the outcomes "regular diploma recipients", "seamless transitioners" and "second year persisters"

```
gen grad = (!mi(chrt_grad) & ontime_grad == 1)
gen seamless_transitioners_any = (enrl_1oct_ninth_yr1_any == 1 & ontime_grad == 1)
gen second_year_persisters = (enrl_1oct_ninth_yr1_any == 1 & enrl_1oct_ninth_yr2_any == 1
& ontime_grad == 1)
```

// Step 4: Create average outcomes by race/ethnicity

```
collapse (mean) grad seamless_transitioners_any second_year_persisters (count) N=sid,
by(race_ethnicity)
```

// Step 5: Format the outcome variables so they read as percentages in the graph

```
foreach var of varlist grad seamless_transitioners_any second_year_persisters {
    replace `var' = (`var' * 100)
    format `var' %9.1f
}
```

// Step 6: Reformat the data file so that one variable contains all the outcomes of interest

// 1. Create 4 observations for each school: ninth grade, hs graduation, seamless college transition and second-year persistence

```
foreach i of numlist 1/4 {
    gen time`i' = `i'
}
```

// 2. Keep only African-American, Asian-American, Hispanic, and White students

```
keep if race_ethnicity == 1 | race_ethnicity == 2 | race_ethnicity == 3 | race_ethnicity
== 5
sort race_ethnicity
gen sortorder = _n
```

// 3. Reshape the data file from wide to long

```
reshape long time , i(sortorder)
```


2. PROGRESSION BY STUDENT RACE/ETHNICITY

// 4. Create a single variable that takes on all the outcomes of interest

```
bysort race_ethnicity: gen outcome = 100 if time == 1
bysort race_ethnicity: replace outcome = grad if time == 2
bysort race_ethnicity: replace outcome = seamless_transitioners_any if time == 3
bysort race_ethnicity: replace outcome = second_year_persisters if time == 4
format outcome %9.1f
```

// Step 7: Prepare to graph the results

// 1. Label the outcome

```
label define outcome 1 "Ninth Graders" 2 "On-time Graduates" ///
3 "Seamless College Transitioners" 4 "Second Year Persisters"
label values time outcome
```

// 2. Generate a cohort label to be used in the footnote for the graph

```
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

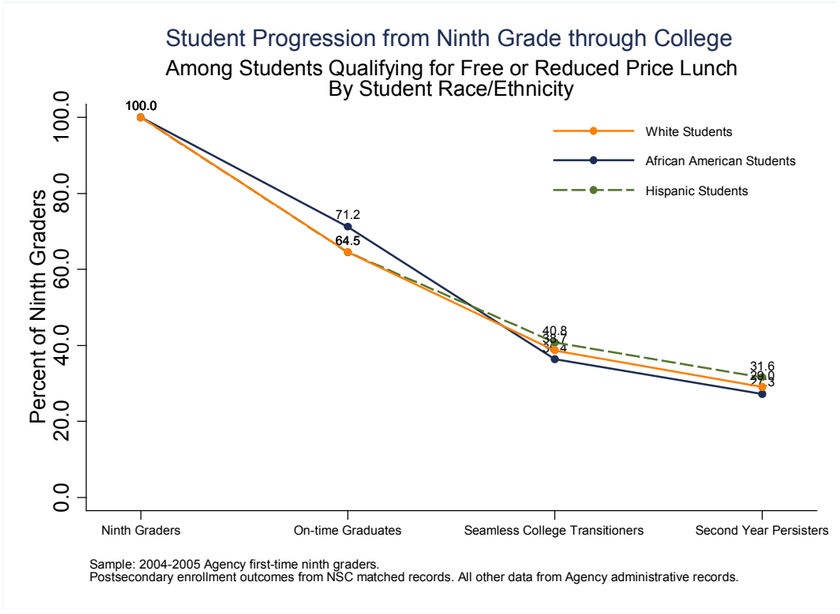
// Step 8: Graph the results

```
#delimit;
twoway (connected outcome time if race_ethnicity==1,
    sort lcolor(dknavy) mlabel(outcome) mlabc(black)mlabs(vsmall) mlabp(12)
    mcolor(dknavy) msymbol(circle) msize(small))
    (connected outcome time if race_ethnicity==2 , sort lcolor(lavender) lpattern(dash)
    mlabel(outcome) mlabs(vsmall) mlabp(12) mlabc(black) mcolor(lavender) msize(small))
    (connected outcome time if race_ethnicity==3 , sort lcolor(dkgreen) lpattern(dash)
    mlabel(outcome) mlabs(vsmall) mlabp(12) mlabc(black) mcolor(dkgreen) msize(small))
    (connected outcome time if race_ethnicity==5 , sort lcolor(orange) mlabel(outcome)
mlabc(black)
    mlabs(vsmall) mlabp(12) mcolor(orange) msymbol(circle) msize(small)),
title("Student Progression from Ninth Grade through College", size(medium))
    subtitle("By Student Race/Ethnicity", size(medsmall))
    xscale(range(.8(.2)4.2))
    xlabel(1 2 3 4 , valuelabels labsize(vsmall))
    ytitle("Percent of Ninth Graders")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
    xtitle("", color(white))
legend(order(2 4 1 3) col(1) position(2) size(vsmall)
    label(1 "African American Students")
    label(2 "Asian American Students")
    label(3 "Hispanic Students")
```

2. PROGRESSION BY STUDENT RACE/ETHNICITY

```
    label(4 "White Students")
    ring(0) region(lpattern(none) lcolor(none) fcolor(none)))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders." "Postsecondary
enrollment outcomes from NSC matched records. All other data from ${agency_name}
administrative records." , size(vsmall));
#delimit cr
graph export "A2_Progression_by_RaceEthnicity.emf", replace
graph save "A2_Progression_by_RaceEthnicity.gph", replace
}
```


3. PROGRESSION BY STUDENT RACE/ETHNICITY, AMONG FRPL-ELIGIBLE STUDENTS



Purpose: This analysis tracks the percent of ninth graders of different races/ethnicities who ever qualified for free or reduce price lunch who complete high school on-time, seamlessly enroll in college, and persist to the second year of college.

Required Analysis File Variables:

sid
race_ethnicity
frpl_ever
chrt_ninth
ontime_grad
enrl_1oct_ninth_yr1_any
enrl_1oct_ninth_yr2_any

Analysis-Specific Sample Restrictions:

- Keep students in ninth grade cohorts for which persistence to the second year of college can be reported.
- Restrict the analysis to include only students who were ever eligible to receive free-or reduced-price lunch throughout their time in your agency, and drop any race/ethnic groups with less than 20 students at any point along the pipeline.

Ask Yourself

- How do differences between races/ethnicities change along the pipeline when only students whoever qualifying for free or reduced price lunch are examined?

3. PROGRESSION BY STUDENT RACE/ETHNICITY, AMONG FRPL-ELIGIBLE STUDENTS

Analytic Technique: Calculate the proportion of first-time ninth graders that progress to each step along the education pipeline.

```
/**** A. Attainment along the Education Pipeline ****/
/**** 3. Progression by Student Race/Ethnicity, Among Frpl-Eligible Students ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe persisting to the second year of college AND are ever FRPL-eligible

```
local chrt_ninth_begin = ${chrt_ninth_begin_persist_yr2}
local chrt_ninth_end = ${chrt_ninth_end_persist_yr2}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end')
keep if frpl_ever == 1
```

// Next, repeat steps 3-9 from the previous analysis

// Step 3: Create variables for the outcomes "regular diploma recipients", "seamless transitioners" and "second year persisters"

```
gen grad = (!mi(chrt_grad) & ontime_grad == 1)
gen seamless_transitioners_any = (enrl_1oct_ninth_yr1_any == 1 & ontime_grad == 1)
gen second_year_persisters = (enrl_1oct_ninth_yr1_any == 1 & enrl_1oct_ninth_yr2_any == 1
& ontime_grad == 1)
```

// Step 4: Create average outcomes by race/ethnicity and drop any race/ethnic groups with fewer than 20 students

```
collapse (mean) grad seamless_transitioners_any second_year_persisters (count) N=sid,
by(race_ethnicity)
drop if N < 20
```

// Step 5: Format the outcome variables so they read as percentages in the graph

```
foreach var of varlist grad seamless_transitioners_any second_year_persisters {
    replace `var' = (`var' * 100)
    format `var' %9.1f
}
```

// Step 6: Reformat the data file so that one variable contains all the outcomes of interest

// 1. Create 4 observations for each school: ninth grade, hs graduation, seamless college transition and second-year persistence

```
foreach i of numlist 1/4 {
    gen time`i' = `i'
}
```

// 2. Keep only African American, Asian American, Hispanic, and White students

```
keep if race_ethnicity == 1 | race_ethnicity == 2 | race_ethnicity == 3 | race_ethnicity
== 5
sort race_ethnicity
gen sortorder = _n
```


3. PROGRESSION BY STUDENT RACE/ETHNICITY, AMONG FRPL-ELIGIBLE STUDENTS

// 3. Reshape the data file from wide to long

```
reshape long time , i(sortorder)
```

// 4. Create a single variable that takes on all the outcomes of interest

```
bysort race_ethnicity: gen outcome = 100 if time == 1
bysort race_ethnicity: replace outcome = grad if time == 2
bysort race_ethnicity: replace outcome = seamless_transitioners_any if time == 3
bysort race_ethnicity: replace outcome = second_year_persisters if time == 4
format outcome %9.1f
```

// Step 7: Prepare to graph the 'results

// 1. Label the outcome

```
label define outcome 1 "Ninth Graders" 2 "On-time Graduates" ///
3 "Seamless College Transitioners" 4 "Second Year Persisters"
label values time outcome
```

// 2. Generate a cohort label to be used in the footnote for the graph

```
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

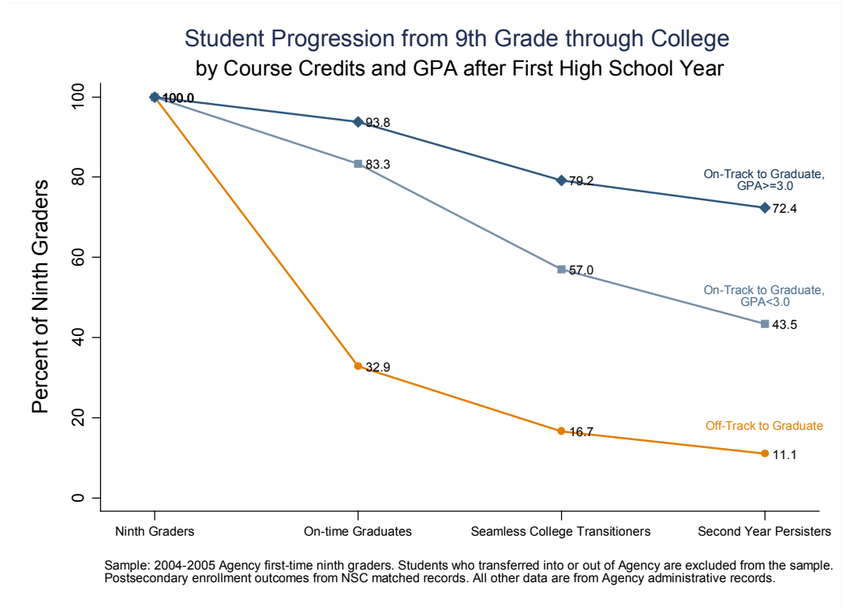
// Step 8: Graph the results

```
#delimit ;
twoway (connected outcome time if race_ethnicity==1 , sort lcolor(dknavy) mlabel(outcome)
    mlabc(black) mlabs(vsmall) mlabp(12) mcolor(dknavy) msymbol(circle) msize(small))
    (connected outcome time if race_ethnicity==3 , sort lcolor(forest_green)
lpattern(dash)
    mlabel(outcome) mlabs(vsmall) mlabp(12) mlabc(black) mcolor(forest_green)
msize(small))
    (connected outcome time if race_ethnicity==5 , sort lcolor(orange) mlabel(outcome)
mlabc(black)
    mlabs(vsmall) mlabp(12) mcolor(orange) msymbol(circle) msize(small)),
title("Student Progression from Ninth Grade through College", size(medium))
    subtitle("Among Students Qualifying for Free or Reduced Price Lunch" "By Student Race/
Ethnicity", size(medsmall))
    xscale(range(.8(.2)4.2))
    xlabel(1 2 3 4, valuelabels labsize(vsmall))
    ytitle("Percent of Ninth Graders")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
    xtitle("", color(white))
```

3. PROGRESSION BY STUDENT RACE/ETHNICITY, AMONG FRPL-ELIGIBLE STUDENTS

```
legend(order(3 1 2) col(1) position(2) size(vsmall)
    label(1 "African American Students")
    label(2 "Hispanic Students")
    label(3 "White Students")
    ring(0) region(lpattern(none) lcolor(none) fcolor(none)))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders." "Postsecondary
enrollment outcomes from NSC matched records. All other data from ${agency_name}
administrative records." , size(vsmall));
#delimit cr
graph export "A3_Progression_by_RaceEthnicity_Frpl.emf", replace
graph save "A3_Progression_by_RaceEthnicity_Frpl.gph", replace
}
```


4. PROGRESSION BY STUDENTS' ON-TRACK STATUS AFTER NINTH GRADE



Purpose: This analysis tracks the percent of ninth graders at different levels of being on-track for graduation who complete high school on-time, seamlessly enroll in college, and then persist to the second year of college.

Required Analysis File Variables:

sid
chrt_ninth
ontrack_sample
ontrack_endyr1*
cum_gpa_yr1*
ontime_grad
enrl_1oct_ninth_yr1_any
enrl_1oct_ninth_yr2_any

Analysis-Specific Sample Restrictions:

- Only include the three most recent ninth grade cohorts for which persistence to second year of college can be reported.
- Restrict the sample to include only students in the on-track analytic sample (students who attended the first semester of ninth grade in the system and never transferred into, or out of the system).
- Students that obtain Special Education diplomas upon high school entry should be excluded from the analytic sample if these students are not required to meet the same graduation requirements as general education students, and if the designation can be made.

Ask Yourself

- How does being on-track for graduation after ninth grade relate to on-time graduation, seamless enrollment, and second year persistence?
- How does being on-track after ninth grade with a higher GPA compare to being on-track with a lower GPA?

4. PROGRESSION BY STUDENTS' ON-TRACK STATUS AFTER NINTH GRADE

Analytic Technique: Calculate the proportion of first-time ninth graders that progressed along the education pipeline.

```
/**** A. Attainment along the Education Pipeline ****/
/**** 4. Progression by Students' On-Track Status After Ninth Grade ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe persisting to the second year of college AND are included in the on-track analysis sample

```
local chrt_ninth_begin = ${chrt_ninth_begin_persist_yr2}
local chrt_ninth_end = ${chrt_ninth_end_persist_yr2}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end')
keep if ontrack_sample == 1
```

// Step 3: Generate on-track indicators that take into account students' GPAs upon completion of their first year in high school

```
label define ot 1 "Off-Track to Graduate" ///
2 "On-Track to Graduate, GPA < 3.0" ///
3 "On-Track to Graduate, GPA >= 3.0", replace
gen ontrack_endyr1_gpa = .
replace ontrack_endyr1_gpa = 1 if ontrack_endyr1 == 0
replace ontrack_endyr1_gpa = 2 if ontrack_endyr1 == 1 & cum_gpa_yr1 < 3 & !mi(cum_gpa_yr1)
replace ontrack_endyr1_gpa = 3 if ontrack_endyr1 == 1 & cum_gpa_yr1 >= 3 & !mi(cum_gpa_
yr1)

assert !mi(ontrack_endyr1_gpa) if !mi(ontrack_endyr1) & !mi(cum_gpa_yr1)
label values ontrack_endyr1_gpa ot
```

// Step 4: Create variables for the outcomes "regular diploma recipients", "seamless transitioners" and "second year persisters"

```
gen grad = (!mi(chrt_grad) & ontime_grad == 1)
gen seamless_transitioners_any = (enrl_1oct_ninth_yr1_any == 1 & ontime_grad == 1)
gen second_year_persisters = (enrl_1oct_ninth_yr1_any == 1 & enrl_1oct_ninth_yr2_any == 1
& ontime_grad == 1)
```

// Step 5: Create average outcomes by on-track status at the end of ninth grade

```
collapse (mean) grad seamless_transitioners_any second_year_persisters (count) N=sid,
by(ontrack_endyr1_gpa)
```

// Step 6: Format the outcome variables so they read as percentages in the graph

```
foreach var of varlist grad seamless_transitioners_any second_year_persisters {
    replace `var' = (`var' * 100)
    format `var' %9.1f
}
```


4. PROGRESSION BY STUDENTS' ON-TRACK STATUS AFTER NINTH GRADE

// Step 7: Reformat the data file so that one variable contains all the outcomes of interest

```
// 1. Create 4 observations for each school: ninth grade, hs graduation, seamless college transition and second-year
persistence
foreach i of numlist 1/4 {
    gen time`i' = `i'
}
// 2. Reshape the data file from wide to long
reshape long time, i(ontrack_endyr1_gpa N)
// 3. Create a single variable that takes on all the outcomes of interest
bysort ontrack_endyr1_gpa: gen outcome = 100 if time == 1
bysort ontrack_endyr1_gpa: replace outcome = grad if time == 2
bysort ontrack_endyr1_gpa: replace outcome = seamless_transitioners_any if time == 3
bysort ontrack_endyr1_gpa: replace outcome = second_year_persisters if time == 4
format outcome %9.1f
```

// Step 8: Prepare to graph the results

```
// 1. Label the outcome
label define outcome 1 "Ninth Graders" 2 "On-time Graduates" ///
3 "Seamless College Transitioners" 4 "Second Year Persisters"
label values time outcome
// 2. Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
// 3. Determine the location of the label for each on-track outcome
sort ontrack_endyr1_gpa _j
foreach obsnum of numlist 4(4)12 {
    local ontrack`obsnum'_label = outcome + 7 in `obsnum'
}
```

// Step 9: Graph the results

```
#delimit ;
twoway (connected outcome time if ontrack_endyr1_gpa == 1,
    sort lcolor(dkorange) mlabel(outcome) mlabc(black) mlabs(vsmall) mlabp(3)
    mcolor(dkorange) msymbol(circle) msize(small))
    (connected outcome time if ontrack_endyr1_gpa == 2, sort lcolor(navy*.6)
    mlabel(outcome) mlabs(vsmall) mlabp(3) mlabc(black) mcolor(navy*.6)
    msymbol(square) msize(small))
    (connected outcome time if ontrack_endyr1_gpa == 3, sort lcolor(navy*.9)
    mlabel(outcome) mlabs(vsmall) mlabp(3) mlabc(black) mcolor(navy*.9)
    msymbol(diamond) msize(small))
```

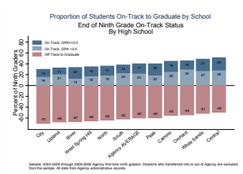
4. PROGRESSION BY STUDENTS' ON-TRACK STATUS AFTER NINTH GRADE

```
    (connected outcome time if ontrack_endyr1_gpa == 4, sort lcolor(navy*.3)
    mlabel(outcome) mlabs(vsmall) mlabp(3) mlabc(black) mcolor(navy*.3)
    msymbol(triangle) msize(small)),
title("Student Progression from 9th Grade through College", size(medium))
    ylabel(, nogrid)
    subtitle("by Course Credits and GPA after First High School Year", size(medsmall))
    xscale(range(.8(.2)4.2)) xlabel(1 2 3 4, valuelabels labsize(vsmall)) xtitle("")
    yscale(range(0(20)100)) ylabel(0(20)100, labsize(small) format(%9.0f))
    ytitle("Percent of Ninth Graders" " ")
text(`ontrack4_label' 4 "Off-Track to Graduate", color(dkorange) size(2))
text(`ontrack8_label' 4 "On-Track to Graduate," "GPA<3.0", color(navy*.8) size(2))
text(`ontrack12_label' 4 "On-Track to Graduate," "GPA>=3.0", color(navy*1.3) size(2))
legend(off)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders. Students who
transferred into or out of ${agency_name} are excluded from the sample." "Postsecondary
enrollment outcomes from NSC matched records. All other data are from ${agency_name}
administrative records.", size(vsmall));
#delimit cr
graph export "A4_Progression_by_OnTrack_Ninth.emf", replace
graph save "A4_Progression_by_OnTrack_Ninth.gph", replace
}
```

B. Ninth to Tenth Grade Transition by On-Track Status

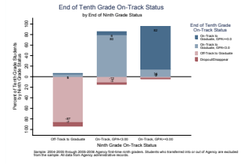
Research suggests that academic performance in ninth grade strongly predicts the likelihood of a student dropping out of high school. In this section, you examine patterns of student retention and on-time transitions from ninth to tenth grade. This information can provide an early warning to an agency with students at-risk of dropping out, and might benefit from targeted support early in their high school careers.

To explore transitions from ninth to tenth grade, use the following model analyses:



1. PROPORTION OF STUDENTS ON-TRACK AT THE END OF NINTH GRADE, BY HIGH SCHOOL

Illustrates what percent of students on-track after ninth grade graduate from each high school and the agency as a whole.

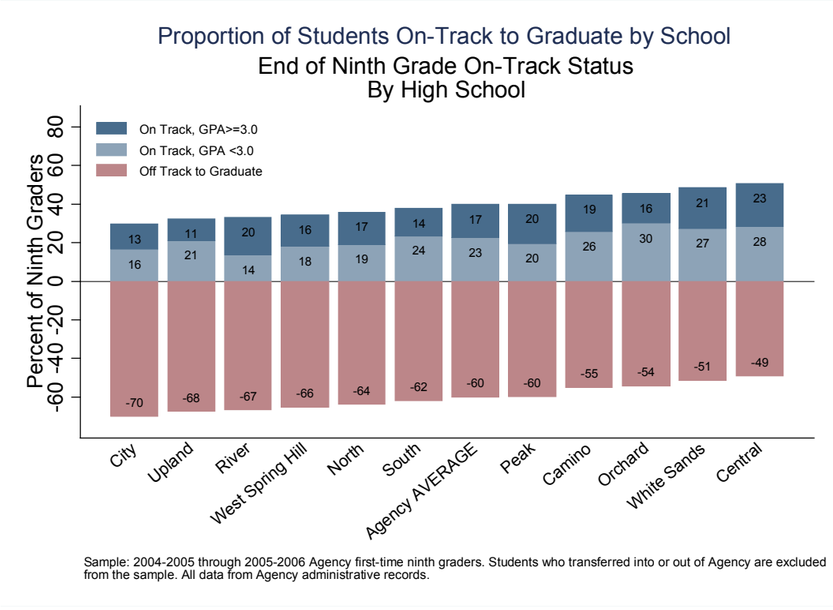


2. NINTH TO TENTH GRADE TRANSITION BY ON-TRACK STATUS

+

Explores how on-track status after ninth grade predicts on-track status in tenth grade.

1. PROPORTION OF STUDENTS ON-TRACK AT THE END OF NINTH GRADE, BY HIGH SCHOOL



Sample: 2004-2005 through 2005-2006 Agency first-time ninth graders. Students who transferred into or out of Agency are excluded from the sample. All data from Agency administrative records.

Purpose: This analysis illustrates what percent of students are on-track after ninth grade graduate from each high school and the agency as a whole. Different levels of on-track for graduation are distinguished by high school.

Required Analysis File Variables:

sid
chrt_ninth
first_hs_name
first_hs_code
ontrack_endyr1*
cum_gpa_yr1*

Analysis-Specific Sample Restrictions: Keep students in ninth grade cohorts you can observe graduating high school on time AND are part of the on-track sample (attended the first semester of ninth grade and never transferred into or out of the system).

Ask Yourself

- How does the percent of students on-track differ by high school (consider the overall height of each bar)?
- How does the percent of students on-track for an advanced versus general diploma differ by high school (consider the different components of each bar)?

Possible Next Steps or Action Plans: Overall school-level results can be disaggregated by student subgroups of interest, (race, FRPL status, and eighth grade academic achievement).

1. PROPORTION OF STUDENTS ON-TRACK AT THE END OF NINTH GRADE, BY HIGH SCHOOL

Analytic Technique: Calculate the proportion of students on-track at each school, and across the agency.

```
/**** B. Ninth to Tenth Grade Transition by On-Track Status ****/
/**** 1. Proportion of Students On-Track at the End of Ninth Grade, By High School ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe graduating high school on time AND are part of the on-track sample (attended the first semester of ninth grade and never transferred into or out of the system)

```
local chrt_ninth_begin = ${chrt_ninth_begin_grad}
local chrt_ninth_end = ${chrt_ninth_end_grad}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end')
keep if ontrack_sample == 1
```

// Step 3: Create on-track categories that account for students' credits earned (already captured in the ontrack_endyr1 variable) and GPA after ninth grade

```
gen ontrack_endyr1_1 = 0
replace ontrack_endyr1_1 = 1 if ontrack_endyr1 == 0
label var ontrack_endyr1_1 "Off-Track to Graduate"

gen ontrack_endyr1_2 = 0
replace ontrack_endyr1_2 = 1 if ontrack_endyr1 ==1 & cum_gpa_yr1 < 3.0
label var ontrack_endyr1_2 "On-Track to Graduate, GPA < 3.0"

gen ontrack_endyr1_3 = 0
replace ontrack_endyr1_3 = 1 if ontrack_endyr1 ==1 & cum_gpa_yr1 >= 3.0 & ///
!mi(cum_gpa_yr1)
label var ontrack_endyr1_3 "On-Track to Graduate, GPA >= 3.0"

assert ontrack_endyr1_1 + ontrack_endyr1_2 + ontrack_endyr1_3 == 1
```

// Step 4: Obtain the agency average for the key variables

```
preserve
    collapse (mean) ontrack_endyr1_? (count) N=sid
    tempfile agency_level
    save `agency_level'
restore
```

// Step 5: Obtain mean rates for each school and append the agency average

```
collapse (mean) ontrack_endyr1_? (count) N=sid, by(first_hs_name first_hs_code)
append using `agency_level'
```

1. PROPORTION OF STUDENTS ON-TRACK AT THE END OF NINTH GRADE, BY HIGH SCHOOL

// Step 6: Provide a hs name label for the appended agency average and shorten hs name

```
replace first_hs_code = 0 if first_hs_code == .
replace first_hs_name = "${agency_name} AVERAGE" if mi(first_hs_name)
replace first_hs_name = subinstr(first_hs_name, " High School", "", .)
```

// Step 7: For students who are off-track upon completion of their first year of high school, convert the values to be negative for ease of visualization in the graph

```
replace ontrack_endyr1_1 = -ontrack_endyr1_1
```

// Step 8: Multiply the average of each outcome by 100 for graphical representation of the rates. Create a variable equal to the sum of the two on-track status variables for easier sorting

```
foreach var of varlist ontrack_endyr1_1 ontrack_endyr1_2 ontrack_endyr1_3  {
replace `var' = (`var' * 100)
format `var' %9.1f
}
gen ontrack_endyr1_sum = ontrack_endyr1_2 + ontrack_endyr1_3
```

// Step 9: Prepare to graph the results

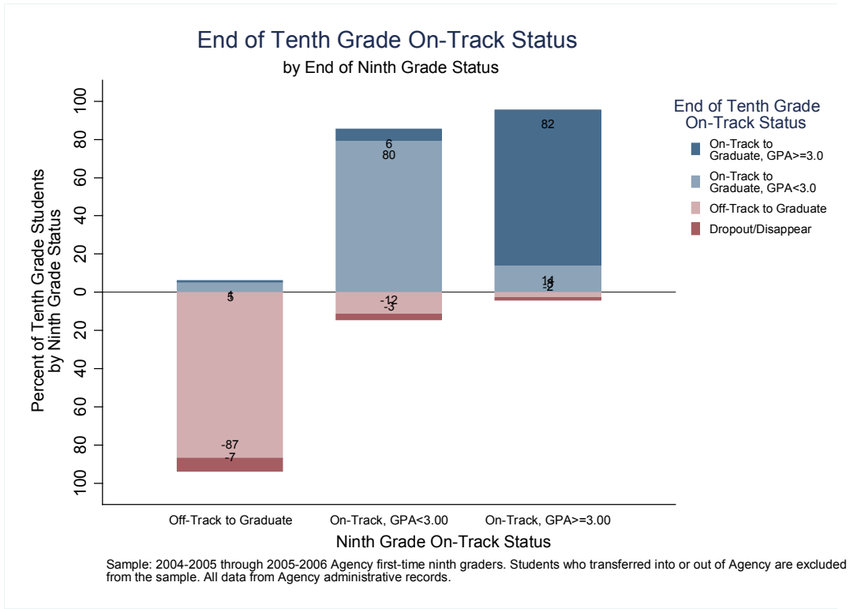
```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```


1. PROPORTION OF STUDENTS ON-TRACK AT THE END OF NINTH GRADE, BY HIGH SCHOOL

// Step 10: Graph the results

```
#delimit ;
graph bar ontrack_endyr1_3 ontrack_endyr1_2 ontrack_endyr1_1,
    over(first_hs_name, gap(20) sort(ontrack_endyr1_sum) label(angle(40)labsize(small)))
    blabel(bar, position(inside) size(2) format(%8.0f))
    bar(3, fcolor(maroon*.6) lcolor(maroon*.6))
    bar(1, fcolor(navy*.5) lcolor(navy*.5))
    bar(2, fcolor(navy*.8) lcolor(navy*.8)) stack
title("Proportion of Students On-Track to Graduate by School", size(medium))
    subtitle("End of Ninth Grade On-Track Status" "By High School")
legend(region(lcolor(white)) position(11) ring(0) order(2 1 3)
    label(3 "Off Track to Graduate")
    label(1 "On Track, GPA <3.0")
    label(2 "On Track, GPA>=3.0")
    symxsize(5) symysize(2) cols(1) size(vsmall))
yline(0, lcolor(black) lwidth(vvthin))
ytitle("Percent of Ninth Graders") yscale(range(-60(20)80)) ylabel(-60(20)80, nogrid)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders. Students who
transferred into or out of ${agency_name} are excluded" "from the sample. All data from
${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "B1_OnTrack_Ninth_by_HS.emf", replace
graph save "B1_OnTrack_Ninth_by_HS.gph", replace
}
```

2. NINTH TO TENTH GRADE TRANSITION BY ON-TRACK STATUS



Sample: 2004-2005 through 2005-2006 Agency first-time ninth graders. Students who transferred into or out of Agency are excluded from the sample. All data from Agency administrative records.

Purpose: This analysis explores how on-track status after ninth grade (the horizontal axis) predicts on-track status in tenth grade (the vertical axis). This analysis is useful for developing early dropout warning indicators for at-risk students as early as the second semester of ninth grade.

Required Analysis File Variables:

sid
chrt_ninth
cum_gpa_yr1*
cum_gpa_yr2*
ontrack_endyr1*
ontrack_endyr2*

Analysis-Specific Sample Restrictions: Keep students in ninth grade cohorts you can observe graduating high school on time AND are part of the on-track sample (attended the first semester of ninth grade and never transferred into or out of the system).

Ask Yourself

- What percent of those in a specific on-track category at the end of ninth grade stay in that same on-track category? For example, what percent of off-track ninth graders continue off-track in tenth grade?
- How might you use an early warning system to help students get back on-track for graduation?

Possible Next Steps or Action Plans: Identify additional risk factors, (chronic absenteeism, prior academic achievement etc.) which can be incorporated into analyses like the one above. This could be used to further understand which students struggle, why they struggle, and interventions to keep them enrolled and engaged.

2. NINTH TO TENTH GRADE TRANSITION BY ON-TRACK STATUS

Analytic Technique: Calculate the proportion of students on-track and off-track to graduate after tenth grade according to their on-track status after ninth grade.

```
/**** B. Ninth to Tenth Grade Transition by On-Track Status ****/
/**** 2. Ninth Grade to Tenth Grade Transition, By On-Track Status ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe graduating high school on time AND are part of the on-track sample

```
local chrt_ninth_begin = ${chrt_ninth_begin_grad}
local chrt_ninth_end = ${chrt_ninth_end_grad}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end')
keep if ontrack_sample == 1
```

// Step 3: Create on-track categories that account for students' credits earned (already captured in the ontrack_endyr1 variable) and GPA after ninth grade

```
gen ontrack_endyr1_gpa = .
replace ontrack_endyr1_gpa = 1 if ontrack_endyr1 == 0
replace ontrack_endyr1_gpa = 2 if ontrack_endyr1 ==1 & cum_gpa_yr1 < 3.0
replace ontrack_endyr1_gpa = 3 if ontrack_endyr1 ==1 & cum_gpa_yr1 >= 3.0 & !mi(cum_gpa_
yr1)
assert !mi(ontrack_endyr1_gpa) if !mi(ontrack_endyr1) & !mi(cum_gpa_yr1)

label define ot 1 "Off-Track to Graduate" 2 "On-Track, GPA<3.00" 3 "On-Track, GPA>=3.00"
label val ontrack_endyr1_gpa ot
```

// Step 4: Create indicators for students upon completion of their second year of high school

```
gen ontrack_endyr2_1 = 0
replace ontrack_endyr2_1 = 1 if ontrack_endyr2 == 0
label var ontrack_endyr2_1 "Off-Track to Graduate Yr2"

gen ontrack_endyr2_2 = 0
replace ontrack_endyr2_2 = 1 if ontrack_endyr2 == 1 & cum_gpa_yr2 < 3.0 & !mi(cum_gpa_yr2)
label var ontrack_endyr2_2 "On-Track, GPA < 3.0"

gen ontrack_endyr2_3 = 0
replace ontrack_endyr2_3 = 1 if ontrack_endyr2 == 1 & cum_gpa_yr2 >= 3.0 & !mi(cum_gpa_
yr2)
label var ontrack_endyr2_3 "On-Track, GPA >= 3.0"

gen ontrack_endyr2_4 = 0
replace ontrack_endyr2_4 = 1 if status_after_yr2==3 | status_after_yr2==4
label var ontrack_endyr2_4 "Dropout/Disappear"

//assert ontrack_endyr2_1 + ontrack_endyr2_2 + ontrack_endyr2_3  + ontrack_endyr2_4 == 1
```

2. NINTH TO TENTH GRADE TRANSITION BY ON-TRACK STATUS

// Step 5: Determine the agency average for each of the indicators created in step 4

```
collapse (mean) ontrack_endyr2_1 ontrack_endyr2_2 ontrack_endyr2_3 ontrack_endyr2_4,
by(ontrack_endyr1_gpa)
foreach var of varlist ontrack_endyr2_1 ontrack_endyr2_2 ontrack_endyr2_3 ontrack_endyr2_4
{
replace `var' = ( `var' * 100)
format `var' %9.1f
}
```

// Step 6: For students who are off-track upon completion of their second year of high school, convert the values to be negative for ease of visualization in the graph.

```
replace ontrack_endyr2_1 = ontrack_endyr2_1 * -1
replace ontrack_endyr2_4 = ontrack_endyr2_4 * -1

// Step 7: Prepare to graph the results
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

// Step 8: Graph the results

```
#delimit ;
graph bar ontrack_endyr2_1 ontrack_endyr2_4 ontrack_endyr2_2 ontrack_endyr2_3 ,
    over(ontrack_endyr1_gpa, label(labsize(vsmall)) gap(50)) outergap(50)
    bar(1, fcolor(maroon*.4) lcolor(maroon*.4))
    bar(2, fcolor(maroon*.8) lcolor(maroon*.8))
    bar(3, fcolor(navy*.5) lcolor(navy*.5))
    bar(4, fcolor(navy*.8) lcolor(navy*.8)) stack
    blabel(bar, size(2) format(%8.0f) position(inside))
legend(symxsize(2) symysize(2) rows(4) size(2)
    region(lcolor(white)) position(2) order(4 3 1 2)
    label(1 "Off-Track to Graduate")
    label(2 "Dropout/Disappear")
    label(3 "On-Track to" "Graduate, GPA<3.0")
    label(4 "On-Track to" "Graduate, GPA>=3.0")
    title("End of Tenth Grade" "On-Track Status", size(small)))
title("End of Tenth Grade On-Track Status", size(medium))
    subtitle("by End of Ninth Grade Status", size(small))
    ytitle("Percent of Tenth Grade Students" "by Ninth Grade Status" " " " ", size(small))
    yscale(range(-100(20)100))
    ylabel(-100(20)100, nogrid labsize(small))
    ylabel(-100 "100" -80 "80" -60 "60" -40 "40" -20 "20" 0 "0" 20 "20" 40 "40" 60 "60"
```


B. Ninth to Tenth Grade Transition

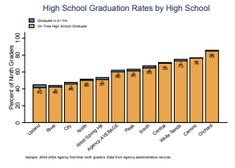
2. NINTH TO TENTH GRADE TRANSITION BY ON-TRACK STATUS

```
    80 "80" 100 "100")
    yline(0, lcolor(black) lwidth(vvthin))
text(-130 60 "Ninth Grade On-Track Status", size(small))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " " " "Sample: `chrt_label' ${agency_name} first-time ninth graders. Students who
transferred into or out of ${agency_name} are excluded" "from the sample. All data from
${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "B2_OnTrack_Tenth_by_OnTrack_Ninth.emf", replace
graph save "B2_OnTrack_Tenth_by_OnTrack_Ninth.gph", replace
}
```

C. High School Graduation

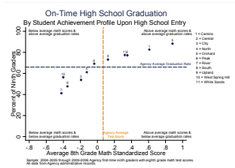
High school graduation is a critical step to higher education. Understanding trends and variations in high school completion rates across schools and student subgroups is essential. These analyses reveal the extent to which high schools may differentially influence student trajectories towards high school completion. After identifying these high schools, you may conduct deeper analyses on your own to explore what drives these outcomes.

To begin exploring high school graduation further, use the analyses below:



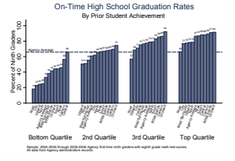
1. HIGH SCHOOL COMPLETION RATES BY SCHOOL

Explores variation in high school completion rates across high schools in the system for both on-time and late high school graduates.



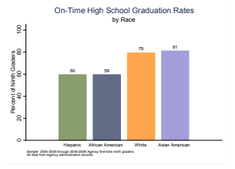
2. HIGH SCHOOL COMPLETION RATES BY AVERAGE 8TH GRADE ACHIEVEMENT

Examines how academic achievement upon high school entry relates to high school completion rates.



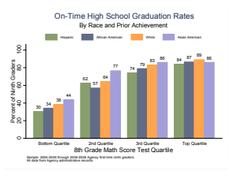
3. HIGH SCHOOL COMPLETION RATES BY 8TH GRADE ACHIEVEMENT QUARTILES

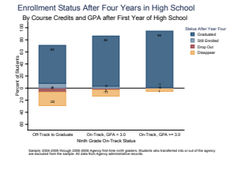
Examines across high schools completion rate variation among students with 8th grade test scores in the same quartile.



4. RACIAL GAPS IN COMPLETION OVERALL AND BY 8TH GRADE ACHIEVEMENT QUARTILES

Displays overall graduation gaps by race. Examines the extent the gap is explained by average differences in academic achievement between racial subgroups at high school entry.

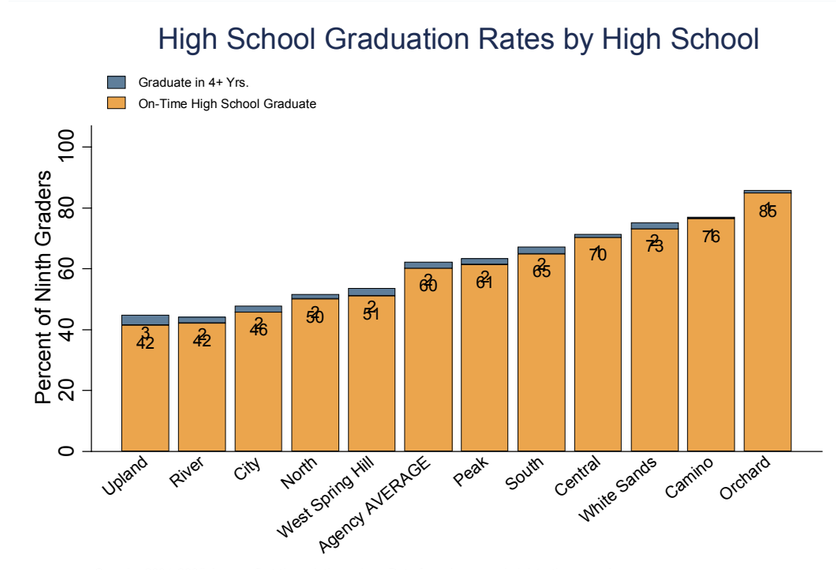




5. ENROLLMENT OUTCOME IN YEAR 4 BY ON-TRACK STATUS AT THE END OF NINTH GRADE

Explores how strongly student performance in ninth grade predicts high school graduation three years later.

1. HIGH SCHOOL COMPLETION RATES BY SCHOOL



Sample: 2004-2005 Agency first-time ninth graders. Data from Agency administrative records.

Purpose: This analysis explores variation in high school completion rates across high schools in the system for both on-time and late high school graduates.

Required Analysis File Variables:

- sid
- chrt_ninth
- hs_diploma
- ontime_grad
- late_grad
- first_hs_code
- first_hs_name

Analysis-Specific Sample Restrictions: Keep students in ninth grade cohorts you can observe graduating high school one year late

Ask Yourself

- Does the ordering of high school completion rates coincide with beliefs key stakeholders have abou these high school?
- Which high schools have the highest and lowest completion rates? Do you know why?

1. HIGH SCHOOL COMPLETION RATES BY SCHOOL

Analytic Technique: Calculate the proportion of students who complete high school by school.

```
/**** C. High School Graduation ****/
/**** 1. High School Graduation Rates by School ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe graduating high school one year late

```
local chrt_ninth_begin = ${chrt_ninth_begin_grad_late}
local chrt_ninth_end = ${chrt_ninth_end_grad_late}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end')
```

// Step 3: Obtain the agency-level high school graduation rates

```
preserve
    collapse (mean) ontime_grad late_grad (count) N = sid
    tempfile agency_level
    save `agency_level'
restore
```

// Step 4: Obtain the school-level high school graduation rates and append the agency average

```
collapse (mean) ontime_grad late_grad (count) N = sid, by(first_hs_name first_hs_code)
append using `agency_level'
```

// Step 5: Provide a hs name label for the appended agency average and shorten hs name

```
replace first_hs_code = 0 if first_hs_code == .
replace first_hs_name = "${agency_name} AVERAGE" if mi(first_hs_name)
replace first_hs_name = subinstr(first_hs_name, " High School", "", .)
```

// Step 6: Multiply the average of each outcome by 100 for graphical representation of the rates

```
foreach var of varlist ontime_grad late_grad {
replace `var' = `var' * 100
format `var' %9.1f
}
```

// Step 7: Prepare to graph the results

// Generate a cohort label to be used in the footnote for the graph

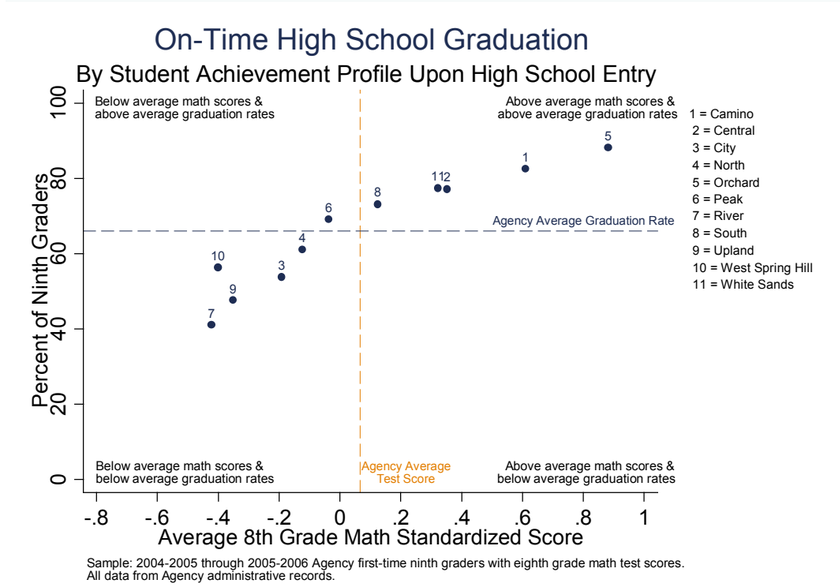
```
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```


1. HIGH SCHOOL COMPLETION RATES BY SCHOOL

// Step 8: Graph the results

```
#delimit ;
graph bar (sum) ontime_grad late_grad, stack over(first_hs_name, label(angle(40)
    labsize(small)) gap(20) sort(ontime_grad))
    blabel(bar, position(inside) color(black) size(small) format(%8.0f))
    bar(1, fcolor(dkorange) fintensity(70) lcolor(black))
    bar(2, fcolor(navy) fintensity(70) lcolor(black))
legend(region(lcolor(white)) symxsize(3) symysize(2) rows(2) order(2 1) size(vsmall)
    position(11) label(1 "On-Time High School Graduate") label(2 "Graduate in 4+ Yrs."))
title("High School Graduation Rates by High School")
    ytitle("Percent of Ninth Graders") yscale(range(0(20)100)) ylabel(0(20)100, nogrid)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders. Data from ${agency_
name} administrative records." , size(vsmall));
#delimit cr
graph export "C1_HS_Grad_by_HS.emf", replace
graph save "C1_HS_Grad_by_HS.gph", replace
}
```

2. HIGH SCHOOL COMPLETION RATES BY AVERAGE 8TH GRADE ACHIEVEMENT



Purpose: This analysis examines the relationship between academic achievement at high school entry and high school completion rates. This analysis is useful to identify high schools that beat the odds. High schools with similar incoming student achievement profiles but different high school graduation rates.

Required Analysis File Variables:

sid
chrt_ninth
test_math_8_std
hs_diploma
first_hs_code
first_hs_name

Analysis-Specific Sample Restrictions:

- Keep students in ninth grade cohorts you can observe graduating high school AND have non-missing eighth grade math scores.
- Drop any high schools with less than 20 students enrolled in ninth grade across the cohorts.

Ask Yourself

- What might explain differences in high school graduation rates for high schools with similar incoming achievement? What might explain differences in incoming achievement for high schools with similar graduation rates?

Possible Next Steps or Action Plans: If substantial variation exists after controlling for average student achievement at high school entry, think about how to share this information across schools. To explore mechanisms that drive school-level differences in high school completion rates, replicate this analysis where the x-axis is a middle school at-risk index (e.g. an index that accounts for whether students failed a core class, were chronically absent, and other information predictive of student achievement in high school) in place of 8th grade test scores.

2. HIGH SCHOOL COMPLETION RATES BY AVERAGE 8TH GRADE ACHIEVEMENT

Analytic Technique: Bivariate scatterplot of school-level average student test scores and high school completion rates.

```
/**** C. High School Graduation ****/
/**** 2. High School Completion Rates by Average 8th Grade Achievement ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe graduating high school AND have non-missing eighth grade math scores

```
local chrt_ninth_begin = ${chrt_ninth_begin_grad}
local chrt_ninth_end = ${chrt_ninth_end_grad}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end') & !mi(test_
math_8_std)
```

// Step 3: Obtain agency-level high school completion rate and prior achievement score for dotted lines. Also get position of their labels

```
summ ontime_grad
local agency_mean_grad = `r(mean)'*100
local agency_mean_grad_label = `agency_mean_grad' + 3
summ test_math_8_std
local agency_mean_test = `r(mean)'
local agency_mean_test_label = `agency_mean_test' + 0.15
```

// Step 4: Obtain school-level high school completion and prior achievement rates

```
collapse (mean) test_math_8_std ontime_grad (count) N = sid, by(first_hs_code first_hs_name)
drop if N < 20
```

// Step 5: Multiply the high school completion rate by 100 for graphical representation of the rates

```
replace ontime_grad = round((ontime_grad * 100), .1)

// Step 6: Shorten high school names and create a legend label for the graph
sort first_hs_name
replace first_hs_name = subinstr(first_hs_name, " High School", "", .)
gen hs_code_label = _n

levelsof first_hs_name, local(hs_names)
local count = 1
local legend_labels ""
foreach hs of local hs_names {
    local legend_labels `"`legend_labels' `count' = `hs'"' `" "'
    local ++count
}
```

2. HIGH SCHOOL COMPLETION RATES BY AVERAGE 8TH GRADE ACHIEVEMENT

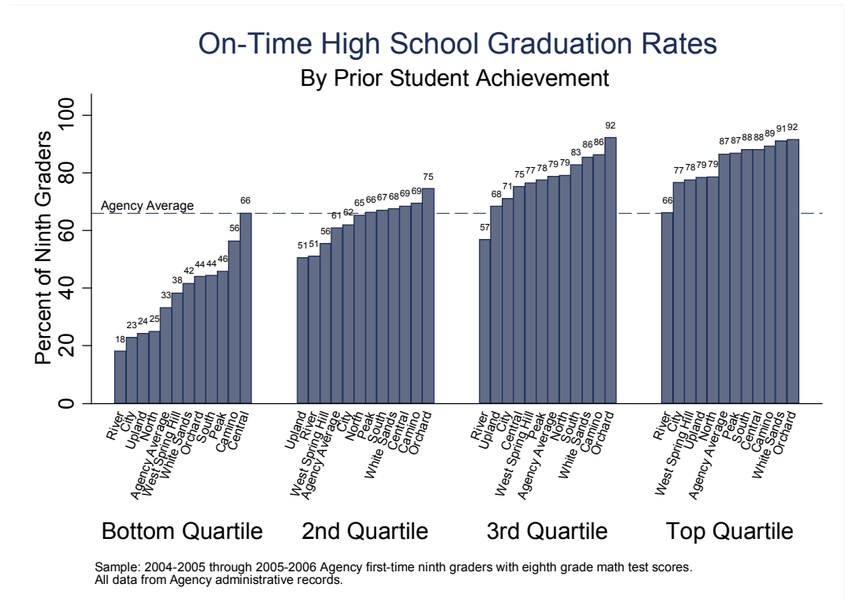
// Step 7: Prepare to graph the results

```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

// Step 8: Graph the results

```
#delimit ;
twoway (scatter ontime_grad test_math_8_std, mlabel(hs_code_label) mlabsize(vsmall)
    mlabposition(12) mlabcolor(dknavy) mstyle(x) msize(small) mcolor(dknavy)),
title("On-Time High School Graduation")
    subtitle("By Student Achievement Profile Upon High School Entry")
    xtitle("Average 8th Grade Math Standardized Score", linegap(0.3))
    ytitle("Percent of Ninth Graders")
    xscale(range(-0.8(0.2)1)) xlabel(-0.8(0.2)1)
    yscale(range(0(20)100)) ylabel(0(20)100, nogrid)
    legend(on order(3) col(1) label(3 `"`legend_labels'"')
    region(color(none)) size(vsmall) position(2) ring(1) linegap(.75))
yline(`agency_mean_grad', lpattern(dash) lcolor(dknavy) lwidth(vvthin))
xline(`agency_mean_test', lpattern(dash) lcolor(dkorange) lwidth(vvthin))
text(`agency_mean_grad_label' .8 "${agency_name} Average Graduation Rate", size(2.0)
color(dknavy))
text(2 `agency_mean_test_label' "${agency_name} Average" "Test Score", size(2.0)
color(dkorange))
text(99 -.5 "Below average math scores &" "above average graduation rates",
    size(vsmall) justification(left))
text(99 0.8 "Above average math scores &" "above average graduation rates",
    size(vsmall) justification(right))
text(2 -0.5 "Below average math scores &" "below average graduation rates",
    size(vsmall) justification(left))
text(2 0.8 "Above average math scores &" "below average graduation rates",
    size(vsmall) justification(right))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' ${agency_name} first-time ninth graders with eighth grade math
test scores." "All data from ${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "C2_HS_Grad_by_Avg_Eighth.emf", replace
graph save "C2_HS_Grad_by_Avg_Eighth.gph", replace
}
```


3. HIGH SCHOOL COMPLETION RATES BY 8TH GRADE ACHIEVEMENT QUARTILES



Purpose: This analysis examines variation in completion rates for high schools among students with 8th grade test scores in the same quartile. The analysis is useful to explore high school completion rates across schools with students in the same quartile or range of achievement. Each high school is repeated as a blue bar in each quartile.

Required Analysis File Variables:

- sid
- chrt_ninth
- hs_diploma
- qrt_8_math
- first_hs_code
- first_hs_name

Analysis-Specific Sample Restrictions:

- Keep students in ninth grade cohorts you can observe graduating high school AND have non-missing eighth grade math scores.
- Drop high schools with less than 20 students in each quartile enrolled in ninth grade across the cohorts.

Ask Yourself

- Looking at the average in each quartile (orange bars), how do 8th grade test scores relate to high school graduation?
- For each quartile of 8th grade test scores (the blue bars), how do graduation rates vary by high school? What is the difference between top and bottom high schools in each quartile?

Possible Next Steps or Action Plans: Highlight comparison schools to show variation across quartiles and explore reasons why students at different schools, but with similar academic profiles at high school entry, are more or less likely to graduate.

3. HIGH SCHOOL COMPLETION RATES BY 8TH GRADE ACHIEVEMENT QUARTILES

Analytic Technique: Calculate the proportion of students, by high school, who complete high school and 8th grade test score quartile for each.

```
/**** C. High School Graduation ****/
/**** 3. High School Completion Rates by 8th Grade Achievement Quartiles ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe graduating high school AND have non-missing eighth grade math scores

```
local chrt_ninth_begin = ${chrt_ninth_begin_grad}
local chrt_ninth_end = ${chrt_ninth_end_grad}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end') & !mi(test_
math_8)
```

// Step 3: Obtain the overall agency-level high school graduation rate along with the position of its label

```
summ ontime_grad
local agency_mean = `r(mean)'*100
local agency_mean_label = `agency_mean'+3
```

// Step 4: Obtain the agency-level high school graduation rates by test score quartile

```
preserve
    collapse (mean) ontime_grad (count) N = sid, by(qrt_8_math)
    tempfile agency_level
    save `agency_level'
restore
```

// Step 5: Obtain school-level high school graduation rates by test score quartile and append the agency-level graduation rates by quartile

```
collapse (mean) ontime_grad (count) N = sid, by(first_hs_code first_hs_name qrt_8_math)
append using `agency_level'
```

// Step 6: Shorten high school names and drop any high schools with fewer than 20 students

```
replace first_hs_code = 0 if first_hs_code == .
replace first_hs_name = "${agency_name} Average" if mi(first_hs_name)
replace first_hs_name = subinstr(first_hs_name, " High School", "", .)
drop if N < 20
```

// Step 7: Multiply the high school completion rate by 100 for graphical representation of the rates

```
replace ontime_grad = round((ontime_grad * 100), .1)
```

// Step 8: Create a variable to sort schools within each test score quartile in ascending order

```
sort qrt_8_math ontime_grad
gen rank = _n
```


3. HIGH SCHOOL COMPLETION RATES BY 8TH GRADE ACHIEVEMENT QUARTILES

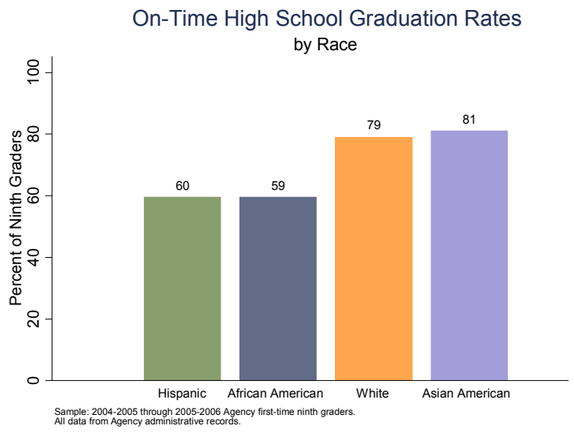
// Step 9: Prepare to graph the results

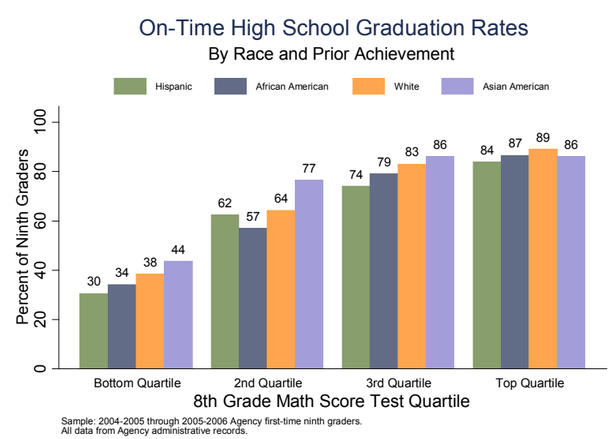
```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

// Step 10: Graph the results

```
#delimit ;
graph bar ontime_grad, over(first_hs_name, sort(rank) gap(0) label(angle(70)
labsize(vsmall)))
    over(qrt_8_math, relabel(1 "Bottom Quartile" 2 "2nd Quartile" 3 "3rd Quartile" 4 "Top
Quartile") gap(400))
    bar(1, fcolor(dknavy) finten(70) lcolor(dknavy) lwidth(thin))
    blabel(bar, format(%8.0f) size(1.5))
    yscale(range(0(20)100)) ylabel(0(20)100, nogrid) legend(off)
title("On-Time High School Graduation Rates")
    subtitle("By Prior Student Achievement", size(msmall))
    ytitle("Percent of Ninth Graders")
    yline(`agency_mean', lpattern(dash) lwidth(vvthin) lcolor(dknavy))
text(`agency_mean_label' 5 "${agency_name} Average", size(vsmall))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders with eighth grade
math test scores." "All data from ${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "C3_HS_Grad_by_Eighth_Qrt.emf", replace
graph save "C3_HS_Grad_by_Eighth_Qrt.gph", replace
}
```

4. RACIAL GAPS IN COMPLETION OVERALL AND BY 8TH GRADE ACHIEVEMENT QUARTILES





Purpose: This analysis displays an overall graduation gap by race, and examines the extent to which this gap is explained by average differences in academic achievement between racial sub-groups at high school entry. The analysis is useful to diagnose whether racial gaps in high school result from persistent academic achievement gaps that emerge in early grades, or if other factors unique to the high school experience drive high school completion rate differences by race.

Required Analysis File Variables:

sid
race_ethnicity
qrt_8_math
ontime_grad
chrt_ninth

Analysis-Specific Sample Restrictions:

- Keep students in ninth grade cohorts you can observe graduating high school AND have non-missing eighth grade math scores.
- Drop any race/ethnic sub-groups with at least 20 students in each quartile (for the second graph). You may further restrict the sample to only include students from the most representative racial/ethnic sub-groups in your agency.

Ask Yourself

- How do racial gaps in graduation rates change after prior achievement is accounted for? Do these gaps change for different prior achievement quartiles?

Possible Next Steps or Action Plans: Repeat analyses for only students that qualify for free or reduced price lunch (FRPL) to explore if racial gaps are better explained by disparities in prior academic achievement and family socioeconomic status.

4. RACIAL GAPS IN COMPLETION OVERALL AND BY 8TH GRADE ACHIEVEMENT QUARTILES

Analytic Technique: Calculate the proportion of students who complete high school by race/ethnicity overall, and by race/ethnicity and 8th grade test score quartile.

```
/**** C. High School Graduation ****/
/**** 4. Graduation Rates by Race Overall and By 8th Grade Achievement Quartiles ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe graduating high school AND have non-missing eighth grade math scores

```
local chrt_ninth_begin = ${chrt_ninth_begin_grad}
local chrt_ninth_end = ${chrt_ninth_end_grad}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end') & !mi(test_
math_8)
```

// Step 3: Obtain the average on-time high school completion rate by race/ethnicity; you will restore in step 8

```
preserve
collapse (mean) ontime_grad (count) N=sid, by(race_ethnicity)
```

// Step 4: Multiply the high school completion rate by 100 for graphical representation of the rates

```
replace ontime_grad = (ontime_grad * 100)
```

// Step 5: Reshape the data wide so that each race is associated with the outcome variable

```
gen id = _n
reshape wide ontime_grad, i(id) j(race_ethnicity)
```

// Step 6: Prepare to graph the results

```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_ninth_begin'-1
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

4. RACIAL GAPS IN COMPLETION OVERALL AND BY 8TH GRADE ACHIEVEMENT QUARTILES

// Step 7: Graph the results (1/2)

```
#delimit ;
graph bar ontime_grad3 ontime_grad1 ontime_grad5 ontime_grad2,
    bargap(25) outergap(100)
    bar(1, fcolor(forest_green*.7) lcolor(forest_green*.7))
    bar(2, fcolor(dknavy*.7) lcolor(dknavy*.7))
    bar(3, fcolor(orange*.7) lcolor(orange*.7))
    bar(4, fcolor(lavender*.85) lcolor(lavender*.85))
    blabel(bar, size(small) format(%8.0f))
text(-4 22 "Hispanic", size(small))
text(-4 40 "African American", size(small))
text(-4 59 "White", size(small))
text(-4 77 "Asian American", size(small))
title("On-Time High School Graduation Rates")
    subtitle("by Race")
    ytitle("Percent of Ninth Graders")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
legend(off)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " " " "Sample: `chrt_label' ${agency_name} first-time ninth graders." "All data from
${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "C4a_HS_Grad_by_Race.emf", replace
graph save "C4a_HS_Grad_by_Race.gph", replace
```

// Step 8: Restore the data and repeat steps 3-6 to obtain completion rates by race/ethnicity and eighth grade test score quartiles. Browse the data to determine if any race/ethnic quartiles have very few students in them.

```
restore
collapse (mean) ontime_grad (count) N=sid, by(race_ethnicity qrt_8_math)
br
replace ontime_grad = (ontime_grad * 100)
reshape wide ontime_grad, i(qrt_8_math N) j(race_ethnicity)
```

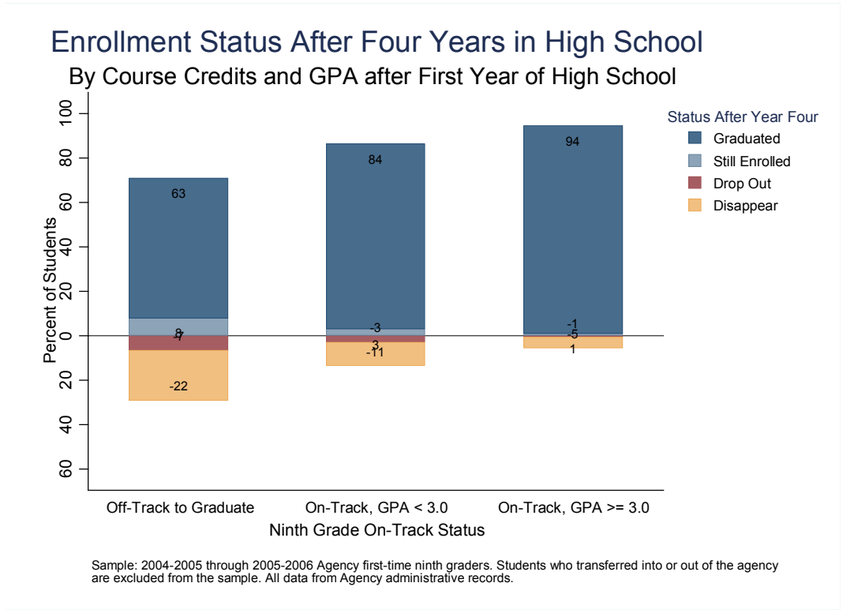
// Step 9: Graph the results (2/2)

```
#delimit ;
graph bar ontime_grad3 ontime_grad1 ontime_grad5 ontime_grad2, over(qrt_8_math,
    relabel(1 "Bottom Quartile" 2 "2nd Quartile" 3 "3rd Quartile" 4 "Top Quartile")
label(labsize(small)))
    bar(1, fcolor(forest_green*.7) lcolor(forest_green*.7)) bar(2, fcolor(dknavy*.7)
lcolor(dknavy*.7))
    bar(3, fcolor(orange*.7) lcolor(orange*.7)) bar(4, fcolor(lavender*.85)
lcolor(lavender*.85))
    blabel(bar, format(%8.0f))
title("On-Time High School Graduation Rates")
    subtitle("By Race and Prior Achievement")
```


4. RACIAL GAPS IN COMPLETION OVERALL AND BY 8TH GRADE ACHIEVEMENT QUARTILES

```
    b1title("8th Grade Math Score Test Quartile")
    ytitle("Percent of Ninth Graders") yscale(range(0(20)100)) ylabel(0(20)100, nogrid)
legend(order(1 2 3 4) row(1) label(1 "Hispanic")
    label(2 "African American") label(3 "White") label(4 "Asian American") size(vsmall)
    symxsize(7) position(inside) ring(1) region(lstyle(none)
    lcolor(none) color(none)))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' ${agency_name} first-time ninth graders." "All data from
${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "C4b_HS_Grad_by_Race_by_Eighth_Qrt.emf", replace
graph save "C4b_HS_Grad_by_Race_by_Eighth_Qrt.gph", replace
}
```

5. ENROLLMENT OUTCOME IN YEAR FOUR BY ON-TRACK STATUS AT THE END OF NINTH GRADE



Purpose: This analysis explores how strongly student performance in ninth grade predicts high school graduation three years later. Building upon our analysis of the relationship between student performance in ninth and tenth grade, the analysis assesses the utility of using course-level performance data early in students' high school careers to assess risk of non-completion, and target students in need of academic and/or socio-emotional support.

Required Analysis File Variables:

sid
chrt_ninth
ontrack_grad_hs_sample*
ontrack_endyr1*
cum_gpa_yr1*
status_after_yr4*

Analysis-Specific Sample Restrictions:

- Keep students in ninth grade cohorts you can observe graduating high school AND are part of the on-track sample (attended the first semester of ninth grade and never transferred into or out of the system).

Ask Yourself

- How does on-track status at the end of ninth grade relate to high school completion status at the end of four years?

Possible Next Steps or Action Plans: Repeat analyses for only students that qualify for free or reduced price lunch (FRPL) to explore whether racial gaps are better explained by disparities in prior academic achievement and family socioeconomic status.

5. ENROLLMENT OUTCOME IN YEAR FOUR BY ON-TRACK STATUS AT THE END OF NINTH GRADE

Analytic Technique: Calculate the proportion of students who graduate high school within four years, dropout, remain enrolled in high school for a fifth year, etc. based on on-track status upon completion of ninth grade.

```
/**** C. High School Graduation ****/
/**** 5. Enrollment Outcome in Year 4 By On-Track Status at the End of Ninth Grade ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in ninth grade cohorts you can observe graduating high school AND have non-missing eighth grade math scores AND are part of the on-track sample

```
local chrt_ninth_begin = ${chrt_ninth_begin_grad}
local chrt_ninth_end = ${chrt_ninth_end_grad}
keep if (chrt_ninth >= `chrt_ninth_begin' & chrt_ninth <= `chrt_ninth_end') & !mi(cum_gpa_
yr1)
keep if ontrack_sample==1
```

// Step 3: Assert that the on-track status after year 4 is not missing

```
label define status 1 "Graduated On-Time" 2 "Still Enrolled" 3 "Dropout" 4 "Disappear",
replace
label values status_after_yr4 status
tab status_after_yr4, m
assert !mi(status_after_yr4)
```

// Step 4: Keep only the variables of interest and generate graduation outcomes after year 4. Assign students as still enrolled if they have a graduation cohort but are not observed to be on-time graduates

```
keep status_after_yr4 ontrack_endyr1 chrt_grad chrt_ninth ontime_grad sid still_enrl
dropout disappear cum_gpa_yr1
gen hs_grad = (status_after_yr4 == 1)
replace still_enrl = 1 if ontime_grad == 0 & !mi(chrt_grad)
```

// Step 5: Ensure that the graduation outcome variables after year 4 are now mutually exclusive for each student

```
assert hs_grad + still_enrl + dropout + disappear == 1
```

// Step 6: Generate on-track indicators that take into account students' GPA upon completion of their first year in high school

```
label define ot 1 "Off-Track to Graduate" 2 "On-Track, GPA < 3.0" ///
3 "On-Track, GPA >= 3.0", replace

gen ontrack_endyr1_gpa = .
replace ontrack_endyr1_gpa = 1 if ontrack_endyr1 == 0
replace ontrack_endyr1_gpa = 2 if ontrack_endyr1 == 1 & cum_gpa_yr1 < 3 & !mi(cum_gpa_yr1)
replace ontrack_endyr1_gpa = 3 if ontrack_endyr1 == 1 & cum_gpa_yr1 >= 3 & !mi(cum_gpa_
yr1)
label values ontrack_endyr1_gpa ot
```

5. ENROLLMENT OUTCOME IN YEAR FOUR BY ON-TRACK STATUS AT THE END OF NINTH GRADE

// Step 7: Create average outcomes by on-track status at the end of ninth grade

```
collapse (mean) hs_grad still_enrl dropout disappear (count) N=sid, by(ontrack_endyr1_gpa)
```

// Step 8: Format the outcome variables so they read as percentages in the graph

```
foreach var of varlist hs_grad still_enrl dropout disappear {
    replace `var' = ( `var' * 100)
    format `var' %9.0f
}
```

// Step 9: For students who dropout or disappear, convert their values to be negative for ease of visualization in the graph

```
foreach var in dropout disappear {
    replace `var' = `var'*-1
}
```

// Step 10: Prepare to graph the results

```
// Generate a cohort label to be used in the footnote for the graph
local temp_end = `chrt_ninth_end'-1
if `chrt_ninth_begin'==`chrt_ninth_end' {
    local chrt_label "`temp_begin'-`chrt_ninth_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_ninth_begin' through `temp_end'-`chrt_ninth_end'"
}
```

// Step 11: Graph the results

```
#delimit ;
graph bar dropout disappear still_enrl hs_grad, over(ontrack_endyr1, gap(100)
label(labsize(2.5)))
    stack blabel(bar, position(inside) color(black) format(%9.0f) size(2.1))
    bar(1, fcolor(maroon*.8) lcolor(maroon*.85))
    bar(2, fcolor(dkorange*.5) lcolor(dkorange*.65) lwidth(vvthin))
    bar(3, fcolor(navy*.5) lcolor(navy*.65) lwidth(vvthin))
    bar(4, fcolor(navy*.8) lcolor(navy*.95) lwidth(vvthin))
legend(col(1) order(4 3 1 2)
    lab(1 "Drop Out")
    lab(2 "Disappear" )
    lab(3 "Still Enrolled")
    lab(4 "Graduated")
    size(2.3) symxsize(2) symysize(2) position(2) region(color(none)) title("Status After
Year Four", size(2.5)))
title("Enrollment Status After Four Years in High School", size(large))
    subtitle("By Course Credits and GPA after First Year of High School", size(medium))
    ytitle("Percent of Students", size(small) margin(2 2 0 0))
    yscale(range(-60(20)100))
    ylabel(-60(20)100, nogrid labsize(small))
    ylabel(-60 "60" -40 "40" -20 "20" 0 "0" 20 "20" 40 "40" 60 "60" 80 "80" 100 "100")
    yline(0, lcolor(black) lwidth(vvthin))
```


C. High School Graduation

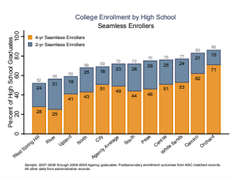
5. ENROLLMENT OUTCOME IN YEAR FOUR BY ON-TRACK STATUS AT THE END OF NINTH GRADE

```
text(-87 50 "Ninth Grade On-Track Status", size(small))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " " " " " "Sample: `chrt_label' ${agency_name} first-time ninth graders. Students
who transferred into or out of the agency"
"are excluded from the sample. All data from ${agency_name} administrative records." ,
size(vsmall));
#delimit cr
graph export "C5_Yr4_Status_by_OnTrack_Ninth.emf", replace
graph save "C5_Yr4_Status_by_OnTrack_Ninth.gph", replace
}
```

D. COLLEGE ENROLLMENT

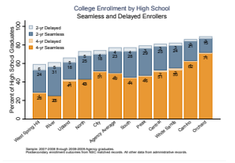
Given the substantial economic and social benefits of a college degree, understanding a high schools' role in preparing students to persist through college is essential. This section provides a series of analyses that highlight college-going rates across high schools in your agency. You will consider whether high school graduates enroll in colleges and universities well-aligned to their incoming academic qualifications. This is one factor that may increase a students' likelihood of college persistence and degree attainment.

To explore these questions, consider the following model analyses:



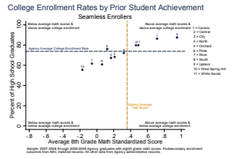
1. COLLEGE ENROLLMENT RATES BY HIGH SCHOOL

Provides an agency snapshot of college enrollment that examines how patterns of college-going by seamless and 2-yr or 4-yr enrollment varies across high schools in the system.



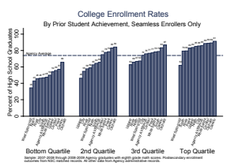
2. SEAMLESS AND DELAYED COLLEGE ENROLLMENT RATES BY HIGH SCHOOL

Provides an agency snapshot of college enrollment that examines how patterns of college-going by seamless or delayed and 2-yr or 4-yr enrollment varies across high schools in the system.



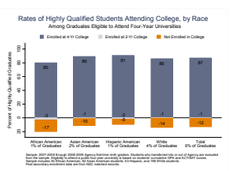
3. COLLEGE ENROLLMENT RATES BY AVERAGE 8TH GRADE ACHIEVEMENT

Examines variation in college enrollment rates for high schools by depicting academic achievement at high school entry. Explains variation in college-going across high schools.



4. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES

Examines variation in college enrollment rates across high schools for students with 8th grade test scores in the same quartile.



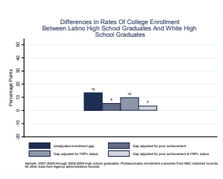
5. RATES OF NON-ENROLLMENT AMONG GRADUATES HIGHLY QUALIFIED TO ATTEND FOUR-YEAR COLLEGES

Examines whether high school graduates enroll in colleges and universities that provide the right academic fit to maximize chances of completion employing the concept of "match."

D. COLLEGE ENROLLMENT

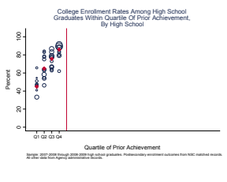
The following three analyses in College Enrollment include the three Strategic Performance Indicators (SPIs) SDP has released to provide deeper insight into the college-going performance of educational systems. These SPIs were conducted using data from a number of SDP's partner agencies. This section of Analyze will help you conduct these analyses yourself. You can read more about the SPIs at **http://www.gse.harvard.edu/sdp/strategic-performance-indicators**.

Strategic Performance Indicators:



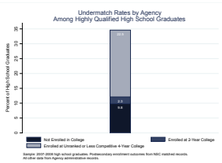
6. GAPS IN RATES OF COLLEGE ENROLLMENT BETWEEN LATINO HIGH SCHOOL GRADUATES AND WHITE SCHOOL GRADUATES

Explores gaps in college enrollment rates by ethnicity, before and after accounting for differences in prior academic achievement, socioeconomic status, and both of these background characteristics.



7. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES - BUBBLES

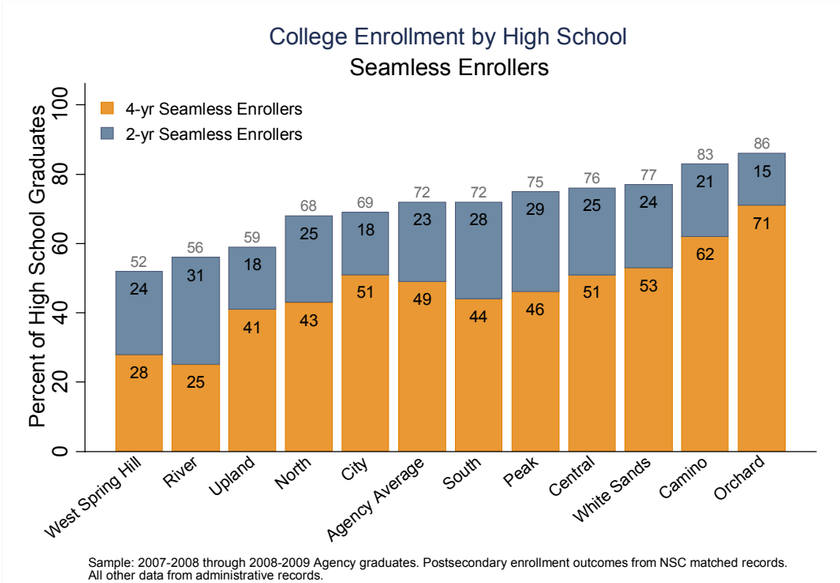
Examines variation in college enrollment rates across high schools for students with 8th grade test scores in the same quartile. This analysis parallels Analysis 4 in this section but provides additional information through the area of each bubble, representing the number of students i each school



8. UNDERMATCH RATES AMONG HIGHLY QUALIFIED HIGH SCHOOL GRADUATES

Examines the prevalence of "undermatch" in the agency—that is, the extent to which high school graduates with strong academic records pursue enrollment in colleges and universities less selective than those for which they are likely qualified. This analysis parallels Analysis 5 in this section but also incorporates information about selectivity of the college attended.

1. COLLEGE ENROLLMENT RATES BY HIGH SCHOOL



Purpose: This analysis provides an agency snapshot of college enrollment to understand how patterns of college-going for high school graduates vary across high schools. By illuminating the extent to which enrollment varies by entry time for seamless enrollers and college level (2- vs. 4-year), the analysis helps diagnose compositional differences for the college-bound population by high school attended.

Required Analysis File Variables:

sid
chrt_grad
enrl_1oct_grad_yr1_any
enrl_1oct_grad_yr1_2yr
enrl_1oct_grad_yr1_4yr
enrl_ever_w2_grad_2yr
enrl_ever_w2_grad_4yr
hs_diploma
last_hs_code
last_hs_name

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation.
- Drop any high schools with less than 20 students in the sample.
- Include only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).

Ask Yourself

- How do college enrollment rates differ by high schools? Why might certain schools have a greater percentage of high school graduates enrolling in college? Do certain schools have higher percentages of 2-year or delayed college enrollers?

Possible Next Steps or Action Plans: Replicate this analysis to include all first-time ninth graders (i.e. ninth grade cohorts) in place of graduates. Additionally, create individual high school reports that provide more details for school administrators (top enrolling institutions of the school's graduates).

1. COLLEGE ENROLLMENT RATES BY HIGH SCHOOL

Analytic Technique: Calculate the proportion of students who enroll in college by high school.

```
/**** D. College Enrollment ****/
/**** 1. Seamless College Enrollment Rates by High School ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
```

// Step 3: Obtain the agency-level average for seamless enrollment

```
preserve
    collapse (sum) enrl_1oct_grad_yr1_2yr enrl_1oct_grad_yr1_4yr hs_diploma
    tempfile agency_level
    save `agency_level'
restore
```

// Step 4: Obtain the school-level averages for seamless enrollment and append on the agency average

```
collapse (sum) enrl_1oct_grad_yr1_2yr enrl_1oct_grad_yr1_4yr hs_diploma, by(last_hs_name
last_hs_code)
append using `agency_level'
```

// Step 5: Provide a hs name label for the appended agency average and shorten hs name

```
replace last_hs_name = "${agency_name} Average" if mi(last_hs_name)
replace last_hs_code = 0 if mi(last_hs_code)
replace last_hs_name = subinstr(last_hs_name, " High School", "", .)
```

// Step 6: Generate percentages of high school grads attending college. Multiply outcomes of interest by 100 for graphical representations of the rates

```
foreach var of varlist enrl_1oct_grad_yr1_* {
    gen pct_`var' = `var' / hs_diploma
    replace pct_`var' = round((pct_`var' * 100))
}
```

// Step 7: Create a total seamless college enrollment rates by summing up the other variables

```
gen total_seamless = pct_enrl_1oct_grad_yr1_2yr + pct_enrl_1oct_grad_yr1_4yr
```

// Step 8: Prepare to graph the results

```
// 1. Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
```

1. COLLEGE ENROLLMENT RATES BY HIGH SCHOOL

```
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

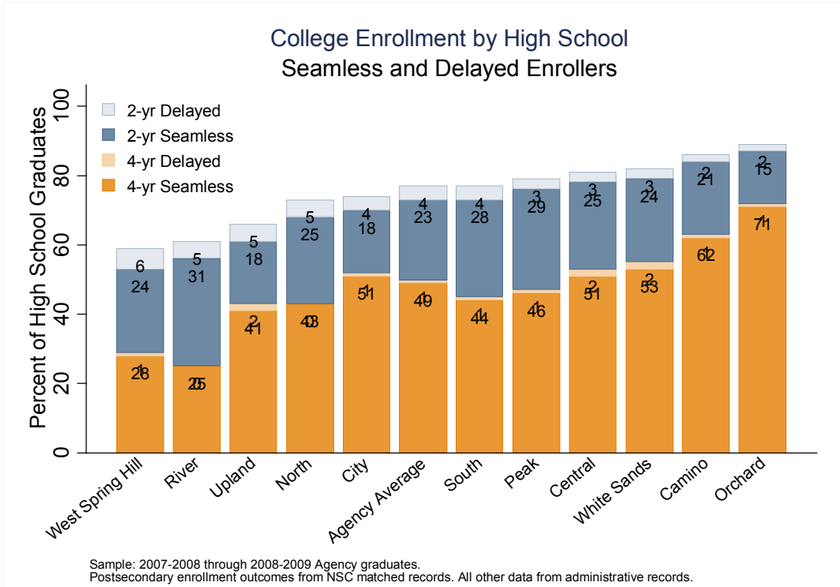
// 2. Generate graphing code to place value labels for the total enrollment rates; change xpos (the position of the first leftmost label) and xposwidth (the horizontal width of the labels) to finetune.

```
sort total_seamless
local total_seamless ""
local num_obs = _N
foreach n of numlist 1/`num_obs' {
    local temp_total_seamless = total_seamless in `n'
    local total_seamless "`total_seamless' `temp_total_seamless'"
}
local total_seamless_label ""
local xpos = 4.8
local xposwidth = 98.7
foreach val of local total_seamless {
    local val_pos = `val' + 3
    local total_seamless_label `"`total_seamless_label' text(`val_pos' `xpos' "`val'",
size(2.5) color(gs7))"'
    local xpos = `xpos' + `xposwidth'/_N
}
disp `"`total_seamless_label'"'
```

// Step 9: Graph the results

```
#delimit ;
graph bar pct_enrl_1oct_grad_yr1_4yr  pct_enrl_1oct_grad_yr1_2yr
    if hs_diploma >= 20, stack over(last_hs_name, label(angle(40) labsize(small)) gap(20)
sort(total_seamless))
    bar(1, fcolor(dkorange) fi(inten80) lcolor(dkorange) lwidth(vvvthin))
    bar(2, fcolor(navy*.8) fi(inten80) lcolor(dknavy*.8) lwidth(vvvthin))
    blabel(bar, position(inside) color(black) size(small))
legend(label(1 "4-yr Seamless Enrollers")
    label(2 "2-yr Seamless Enrollers")
    position(11) ring(0) symxsize(2) symysize(2) rows(2) size(small) region(lstyle(none)
lcolor(none) color(none)))
title("College Enrollment by High School", size(medium))
    ytitle("Percent of High School Graduates")
    subtitle("Seamless Enrollers")
    `total_seamless_label'
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' ${agency_name} graduates. Postsecondary enrollment outcomes
from NSC matched records." "All other data from administrative records.", size(vsmall));
#delimit cr
graph export "D1_Col_Enrl_Seamless_by_HS.emf", replace
graph save "D1_Col_Enrl_Seamless_by_HS.gph", replace
}
```


2. SEAMLESS AND DELAYED COLLEGE ENROLLMENT RATES BY HIGH SCHOOL



Purpose: This analysis provides an agency snapshot of college enrollment to understand how patterns of college-going for high school graduates vary across high schools. By illuminating the extent to which enrollment varies by entry time (seamless vs. delayed) and college level (2- vs. 4-year), the analysis helps diagnose compositional differences for the college-bound population by high school attended.

Required Analysis File Variables:

- sid
- chrt_grad
- enrl_1oct_grad_yr1_any
- enrl_1oct_grad_yr1_2yr
- enrl_1oct_grad_yr1_4yr
- enrl_ever_w2_grad_2yr
- enrl_ever_w2_grad_4yr
- hs_diploma
- last_hs_code
- last_hs_name

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation.
- Drop any high schools with less than 20 students in the sample.
- Include only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).

Ask Yourself

- How do college enrollment rates differ by high schools? Why might certain schools have a greater percentage of high school graduates enrolling in college? Do certain schools have higher percentages of 2-year or delayed college enrollers?

Possible Next Steps or Action Plans: Replicate this analysis to include all first-time ninth graders (i.e. ninth grade cohorts) in place of graduates. Additionally, create individual high school reports that provide more details for school administrators (top enrolling institutions of the school's graduates).

2. SEAMLESS AND DELAYED COLLEGE ENROLLMENT RATES BY HIGH SCHOOL

Analytic Technique: Calculate the proportion of graduates who enroll in four-year institutions across high schools according to the selectivity ranking of the postsecondary institutions attended.

```
/**** D. College Enrollment ****/
/**** 2. Seamless and Delayed College Enrollment Rates by High School ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
```

// Step 3: Create binary outcomes for late enrollers

```
gen late_any = enrl_1oct_grad_yr1_any==0 &  enrl_ever_w2_grad_any==1
gen late_4yr = enrl_1oct_grad_yr1_any==0 &  enrl_ever_w2_grad_4yr==1
gen late_2yr = enrl_1oct_grad_yr1_any==0 &  enrl_ever_w2_grad_2yr==1
assert late_4yr +  late_2yr == late_any
```

// Step 4: Obtain the agency average for seamless and delayed enrollment

```
preserve
    collapse (sum) enrl_1oct_grad_yr1_2yr enrl_1oct_grad_yr1_4yr late_4yr late_2yr hs_
diploma
    tempfile agency_level
    save `agency_level'
restore
```

// Step 4: Obtain the school-level averages for seamless and delayed enrollment and append on the agency average

```
collapse (sum) enrl_1oct_grad_yr1_2yr enrl_1oct_grad_yr1_4yr late_4yr late_2yr hs_diploma,
by(last_hs_name last_hs_code)
append using `agency_level'
```

// Step 5: Provide a hs name label for the appended agency average and shorten hs name

```
replace last_hs_name = "${agency_name} Average" if mi(last_hs_name)
replace last_hs_code = 0 if mi(last_hs_code)
replace last_hs_name = subinstr(last_hs_name, " High School", "", .)
```

// Step 6: Generate percentages of high school grads attending college. Multiply outcomes of interest by 100 for graphical representations of the rates

```
foreach var of varlist enrl_1oct_grad_yr1_* late_* {
    gen pct_`var' = `var' / hs_diploma
    replace pct_`var' = round((pct_`var' * 100))
}
```


2. SEAMLESS AND DELAYED COLLEGE ENROLLMENT RATES BY HIGH SCHOOL

// Step 7: Create total college enrollment rates by summing up the other variables; you can add additional labels as you see fit

```
gen total = pct_enrl_1oct_grad_yr1_2yr + pct_enrl_1oct_grad_yr1_4yr + pct_late_4yr + pct_
late_2yr
gen total_seamless = pct_enrl_1oct_grad_yr1_2yr + pct_enrl_1oct_grad_yr1_4yr
```

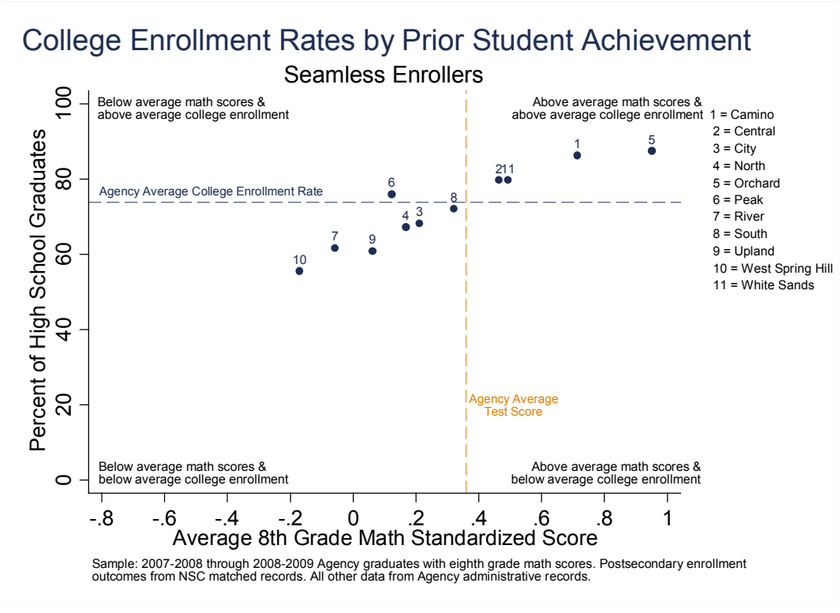
// Step 8: Prepare to graph the results

```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

// Step 9: Graph the results

```
#delimit ;
graph bar pct_enrl_1oct_grad_yr1_4yr pct_late_4yr pct_enrl_1oct_grad_yr1_2yr pct_late_2yr
    if hs_diploma >= 20, over(last_hs_name, label(angle(40)labsize(small)) gap(20)
sort(total))
    bar(1, fcolor(dkorange) fi(inten80) lcolor(dkorange) lwidth(vvvthin))
    bar(2, fcolor(dkorange*.4) fi(inten80) lcolor(dkorange*.4) lwidth(vvvthin))
    bar(3, fcolor(navy*.8) fi(inten80) lcolor(navy*.8) lwidth(vvvthin))
    bar(4, fcolor(navy*.4) fi(inten30) lcolor(navy*.4) lwidth(vvvthin)) stack
    blabel(bar, position(inside) color(black) size(small))
legend(label(1 "4-yr Seamless")
    label(2 "4-yr Delayed")
    label(3 "2-yr Seamless")
    label(4 "2-yr Delayed")
    position(11) order(4 3 2 1) ring(0) symxsize(2) symysize(2) rows(4) size(small)
region(lstyle(none) lcolor(none) color(none)))
title("College Enrollment by High School", size(medium))
    ytitle("Percent of High School Graduates")
    subtitle("Seamless and Delayed Enrollers")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' ${agency_name} graduates."
"Postsecondary enrollment outcomes from NSC matched records. All other data from
administrative records.", size(vsmall));
#delimit cr
graph export "D2_Col_Enrl_Seamless_Delayed_by_HS.emf", replace
graph save "D2_Col_Enrl_Seamless_Delayed_by_HS.gph", replace
}
```

3. COLLEGE ENROLLMENT RATES BY AVERAGE 8TH GRADE ACHIEVEMENT



Sample: 2007-2008 through 2008-2009 Agency graduates with eighth grade math scores. Postsecondary enrollment outcomes from NSC matched records. All other data from Agency administrative records.

Purpose: This analysis displays variations in college enrollment rates across high schools by examining the extent to which academic achievement at high school entry explains variation in college-going across high schools. This analysis is useful to identify high schools with similar incoming student achievement profiles but divergent college enrollment rates; or on the other hand, high schools with similar college-going rates but different academic performance at high school entry.

Required Analysis File Variables:

sid
chrt_grad
enrl_1oct_grad_yr1_any
test_math_8_std
last_hs_code
last_hs_name

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND have non-missing eighth grade test scores.
- Include only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).

Ask Yourself

- What might explain variation in college enrollment rates for high schools with similar incoming achievement? What might explain variation in incoming achievement for high schools with similar college enrollment rates?

3. COLLEGE ENROLLMENT RATES BY AVERAGE 8TH GRADE ACHIEVEMENT

Possible Next Steps or Action Plans: Repeat this analysis to include all first-time ninth graders (i.e. ninth grade cohorts) in place of graduates, and explore college enrollment within two years of high school completion. Additionally, replicate this analysis to explore the relationship between college enrollment and students' ELA achievement at high school entry. Consider why schools with similar incoming student profiles may report dramatically different college-going rates. Conversely, consider why schools with disimilar student bodies report similar matriculation rates to college.

3. COLLEGE ENROLLMENT RATES BY AVERAGE 8TH GRADE ACHIEVEMENT

Analytic Technique: Bivariate scatterplot of school-level average student test scores and college enrollment rates.

```
/**** D. College Enrollment ****/
/**** 3. College Enrollment Rates by Average 8th Grade Achievement ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND have non-missing eighth grade math scores

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end') & !mi(test_math_8_
std)
```

// Step 3: Obtain agency-level college enrollment rate and prior achievement score for dotted lines. Also get position of their labels

```
summ enrl_1oct_grad_yr1_any
local agency_mean_enroll = `r(mean)'*100
local agency_mean_enroll_label = `agency_mean_enroll' + 3
summ test_math_8_std
local agency_mean_test = `r(mean)'
local agency_mean_test_label = `agency_mean_test' + 0.15
```

// Step 4: Obtain school-level college enrollment rates and prior achievement scores

```
collapse (mean) test_math_8_std enrl_1oct_grad_yr1_any (count) N = sid, by(last_hs_code
last_hs_name)
```

// Step 5: Multiply the college enrollment rate by 100 for graphical representation of the rates

```
replace enrl_1oct_grad_yr1_any = round((enrl_1oct_grad_yr1_any * 100), .1)
```

// Step 6: Shorten high school names and create a legend label for the graph

```
sort last_hs_name
replace last_hs_name = subinstr(last_hs_name, " High School", "", .)
gen hs_code_label = _n

levelsof last_hs_name, local(hs_names)
local count = 1
local legend_labels ""
foreach hs of local hs_names {
    local legend_labels `"`legend_labels' `count' = `hs'"' `" "'
    local ++count
}
```


3. COLLEGE ENROLLMENT RATES BY AVERAGE 8TH GRADE ACHIEVEMENT

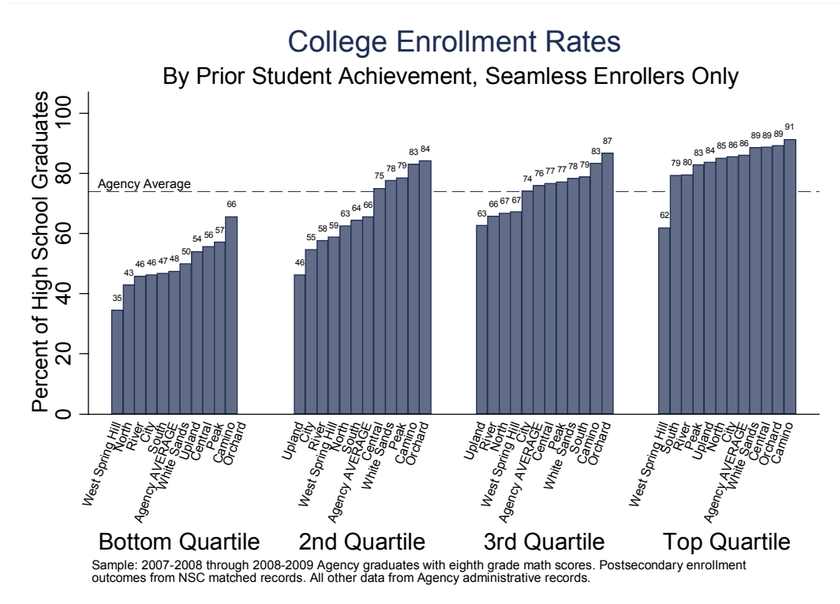
// Step 7: Prepare to graph the results

```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

// Step 8: Graph the results

```
#delimit ;
twoway (scatter enrl_1oct_grad_yr1_any test_math_8_std, mlabel(hs_code_label)
mlabsize(vsmall)
    mlabposition(12) mlabcolor(dknavy) mstyle(x) msize(small) mcolor(dknavy)),
title("College Enrollment Rates by Prior Student Achievement")
    subtitle("Seamless Enrollers")
    xtitle("Average 8th Grade Math Standardized Score", linegap(0.3))
    ytitle("Percent of High School Graduates" " ")
xscale(range(-0.8(0.2)1)) xlabel(-0.8(0.2)1)
yscale(range(0(20)100)) ylabel(0(20)100, nogrid)
legend(on order(3) col(1) label(3 `"`legend_labels'"')
    region(color(none)) size(vsmall) position(2) ring(1) linegap(.75))
    yline(`agency_mean_enroll', lpattern(dash) lcolor(dknavy) lwidth(vvthin))
    xline(`agency_mean_test', lpattern(dash) lcolor(dkorange) lwidth(vvthin))
text(`agency_mean_enroll_label' -0.45 "${agency_name} Average College Enrollment Rate",
size(2.0) color(dknavy))
text(20 `agency_mean_test_label' "${agency_name} Average" "Test Score", size(2.0)
color(dkorange))
text(99 -0.5 "Below average math scores &" "above average college enrollment",
    size(vsmall) justification(left))
text(99 0.8 "Above average math scores &" "above average college enrollment",
    size(vsmall) justification(right))
text(2 -0.5 "Below average math scores &" "below average college enrollment",
    size(vsmall) justification(left))
text(2 0.8 "Above average math scores &" "below average college enrollment",
    size(vsmall) justification(right))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' ${agency_name} graduates with eighth grade math scores.
Postsecondary enrollment"
"outcomes from NSC matched records. All other data from ${agency_name} administrative
records.", size(vsmall));
#delimit cr
graph export "D3_Col_Enrl_by_Avg_Eighth.emf", replace
graph save "D3_Col_Enrl_by_Avg_Eighth.gph", replace
}
```

4. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES



Purpose: This analysis explores whether variation in college enrollment across high schools is similar among low-, middle, and top-achieving students. It also considers whether overall variation across schools derives from concentrated divergence among students scoring in a particular achievement range. Additionally, the analysis facilitates granular school-to-school comparisons to identify those especially under-, or over-performing within each achievement quartile. Finally, the analysis also helps identify which student subgroups require additional resources and support within each school.

Required Analysis File Variables:

sid
chrt_grad
enrl_1oct_grad_yr1_any
qrt_8_math_std
last_hs_code
last_hs_name

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND have non-missing eighth grade test scores.
- Drop high schools with less than 20 students in each quartile enrolled in ninth grade across the cohorts.
- Keep only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).

Ask Yourself

- After looking at the average in each quartile (the orange bars), how do 8th grade test scores relate to college enrollment? Within each quartile of 8th grade test scores (the blue bars), how do enrollment rates vary by high school? What is the difference between top and bottom performing high schools in each quartile?

4. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES

Possible Next Steps or Action Plans: Repeat this analysis to include all first-time ninth graders (i.e. ninth grade cohorts) in place of graduates, and explore college enrollment within two years of high school completion. Additionally, replicate this analysis to explore the relationship between college enrollment and students' ELA achievement at high school entry. Consider why schools with similar incoming student profiles may report dramatically different college-going rates. Conversely, consider why schools with distinct student bodies may report similar matriculation rates to college.

4. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES

Analytic Technique: Calculate the proportion of graduates who enrolled in college by October 1st following their high school graduation year by high school and 8th grade test score quartile.

```
/**** D. College Enrollment ****/
/**** 4. College Enrollment Rates by 8th Grade Achievement Quartiles ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND have non-missing eighth grade math scores

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end') & !mi(qrt_8_math)
```

// Step 3: Obtain the overall agency-level high school graduation rate for dotted line along with the position of its label

```
summ enrl_1oct_grad_yr1_any
local agency_mean = `r(mean)'*100
local agency_mean_label = `agency_mean'+3
```

// Step 4: Obtain the agency-level college enrollment rate by test score quartile

```
preserve
    collapse (mean) enrl_1oct_grad_yr1_any (count) N = sid, by(qrt_8_math)
    tempfile agency_level
    save `agency_level'
restore
```

// Step 5: Obtain school-level college enrollment rates by test score quartile and append the agency-level enrollment rates by quartile

```
collapse (mean) enrl_1oct_grad_yr1_any (count) N = sid, by(last_hs_code last_hs_name
qrt_8_math)
append using `agency_level'
```

// Step 6: Shorten high school names and drop any high schools with fewer than 20 students

```
replace last_hs_code = 0 if last_hs_code == .
replace last_hs_name = "${agency_name} AVERAGE" if mi(last_hs_name)
replace last_hs_name = subinstr(last_hs_name, " High School", "", .)
drop if N < 20
```

// Step 7: Multiply the college enrollment rate by 100 for graphical representation of the rates

```
replace enrl_1oct_grad_yr1_any = round((enrl_1oct_grad_yr1_any * 100), .1)
```


4. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES

// Step 8: Create a variable to sort schools within each test score quartile in ascending order

```
sort qrt_8_math enrl_1oct_grad_yr1_any
gen rank = _n
```

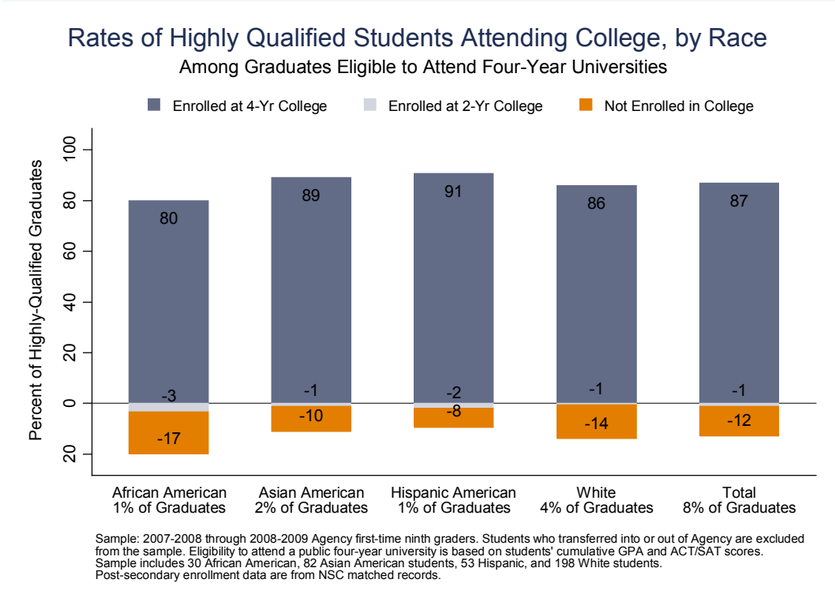
// Step 9: Prepare to graph the results

```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

// Step 10: Graph the results

```
#delimit ;
graph bar enrl_1oct_grad_yr1_any, over(last_hs_name, sort(rank) gap(0) label(angle(70)
labsize(vsmall)))
    over(qrt_8_math, relabel(1 "Bottom Quartile" 2 "2nd Quartile" 3 "3rd Quartile" 4 "Top
Quartile") gap(400))
    bar(1, fcolor(dknavy) finten(70) lcolor(dknavy) lwidth(thin))
    blabel(bar, position(outside) format(%8.0f) size(tiny))
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
legend(off)
title("College Enrollment Rates")
    subtitle("By Prior Student Achievement, Seamless Enrollers Only", size(msmall))
    ytitle("Percent of High School Graduates")
    yline(`agency_mean', lpattern(dash) lwidth(vvthin) lcolor(dknavy))
text(`agency_mean_label' 5 "${agency_name} Average", size(vsmall))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' ${agency_name} graduates with eighth grade math scores.
Postsecondary enrollment" "outcomes from NSC matched records. All other data from
${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "D4_Col_Enrl_by_Eighth_Qrt.emf", replace
graph save "D4_Col_Enrl_by_Eighth_Qrt.gph", replace
}
```

5. RATES OF COLLEGE ENROLLMENT AMONG GRADUATES HIGHLY QUALIFIED TO ATTEND FOUR-YEAR COLLEGES, BY COLLEGE TYPE



Purpose: Research consistently finds wide variation in rates of persistence and completion across postsecondary institutions. This analysis examines whether high school graduates enroll in colleges and universities that provide the right academic fit to maximize their chances of completion. "Match"describes the extent high school graduates with strong academic records attend colleges and universities that allow them to take advantage of their ambition and abilities. While "matching" to an appropriately selective college is only one factor to consider when choosing a postsecondary institution, the implications of under-matching (i.e. lower rates of persistence and degree completion) suggest students should be encouraged to attend realistic, yet challenging postsecondary institutions.

Required Analysis File Variables:

sid
race_ethnicity
chrt_grad
highly_qualified
enrl_1oct_grad_yr1_any
enrl_1oct_grad_yr1_4yr
enrl_1oct_grad_yr1_2yr

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation.
- Include only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).
- Include only highly qualified high school graduates (i.e. students who have obtained a high school diploma on time with 1) a cumulative GPA of 3.0 or higher and Math/Verbal SAT score of 1300 or higher, or 2) a cumulative GPA of 3.3 or higher and Math/Verbal SAT score of 1200 or higher, or 3) a cumulative GPA of 3.7 or higher and Math/Verbal SAT score of at least 1100).
- Drop race/ethnic groups with less than 20 students eligible to attend a four-year university.

5. RATES OF COLLEGE ENROLLMENT AMONG GRADUATES HIGHLY QUALIFIED TO ATTEND FOUR-YEAR COLLEGES, BY COLLEGE TYPE

Ask Yourself

- Among highly qualified students, which race/ethnicities seem to face the greatest undermatch rates?

Possible Next Steps or Action Plans: This analysis leads to important questions that warrant further exploration. What factors drive undermatch differences across student subgroups and high schools? To what extent is undermatching concentrated among first-time college-goers? To what extent is undermatching driven by students' proximity to 2-year versus 4-year institutions? What college aspirations do incoming ninth graders hold, and do these aspirations change by the time they enter or complete 12th grade? To what extent are teachers, counselors, and administrators supported to work with students to cultivate postsecondary aspirations and weigh factors in the college selection process?

5. RATES OF COLLEGE ENROLLMENT AMONG GRADUATES HIGHLY QUALIFIED TO ATTEND FOUR-YEAR COLLEGES, BY COLLEGE TYPE

Analytic Technique: Calculate the proportion of highly qualified graduates who do not enroll in college, enroll in 2-year college, and enroll in least competitive and unranked 4-year colleges the fall following high school graduation.

```
/**** D. College Enrollment ****/
/**** 5. Rates of College Enrollment Among Graduates Highly Qualified to Attend Four-Year
Colleges, By College Type ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
```

// Step 3: Get total number of students in sample

```
gen total_count = _N
```

// Step 4: Further restrict sample to include only highly qualified students

```
keep if highly_qualified == 1
```

// Step 5: Create "undermatch" outcomes

```
gen no_college = (enrl_1oct_grad_yr1_any == 0)
gen enrl_2yr = (enrl_1oct_grad_yr1_2yr == 1)
gen enrl_4yr = (enrl_1oct_grad_yr1_4yr == 1)
```

// Step 4: Create agency-level outcomes for total undermatching rates

```
preserve
    collapse (mean) no_college enrl_2yr enrl_4yr total_count (count) N = sid
    gen group = 5
    tempfile total
    save `total'
restore
```

// Step 5: Create race/ethnicity-level outcomes for undermatching rates by race/ethnicity

```
collapse (mean) no_college enrl_2yr enrl_4yr total_count (count) N = sid , by(race_
ethnicity)
append using `total'

replace group = 1 if race_ethnicity==1
replace group = 2 if race_ethnicity==2
replace group = 3 if race_ethnicity==3
replace group = 4 if race_ethnicity==5
drop if mi(group)
```


5. RATES OF COLLEGE ENROLLMENT AMONG GRADUATES HIGHLY QUALIFIED TO ATTEND FOUR-YEAR COLLEGES, BY COLLEGE TYPE

// Step 7: Multiply the college enrollment rate by 100 for graphical representation of the rates

```
foreach v of varlist no_college enrl_2yr enrl_4yr {
    replace `v' = round(`v'*100, .1)
}
```

// Step 8: Multiply the outcome variables corresponding to undermatching by "-1" to visually display these rates as negative values

```
foreach var of varlist no_college enrl_2yr {
    replace `var' = `var'*-1
}
```

// Step 9: Prepare to graph the results

```
// 1. Create labels for numbers in graph
gen pct_total = N/total_count
sort group
local numobs = _N
foreach v of numlist 1/`numobs' {
    local pct_`v' = round(pct_total*100) in `v'
    local count_`v' = N in `v'
}
// 2. Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

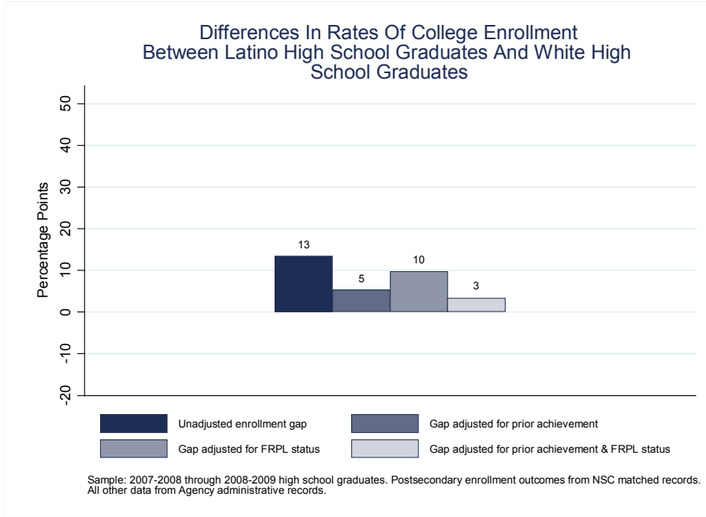
// Step 10: Graph the results

```
#delimit ;
graph bar enrl_4yr enrl_2yr no_college, stack over(group,
    relabel(1 `""African American" "`pct_1'% of Graduates""' 2 `""Asian American"
"`pct_2'% of Graduates""' 3 `""Hispanic American" "`pct_3'% of Graduates""' 4 `""White"
"`pct_4'% of Graduates""' 5 `""Total" "`pct_5'% of Graduates""')
    label(labsize(2.5)) gap(80)) blabel(bar, format(%9.0f) size(small) position(inside)
color(black))
    bar(1, fcolor(dknavy*.7) lcolor(dknavy*.7) lwidth(vvthin))
    bar(2, fcolor(dknavy*.2) lcolor(dknavy*.2) lwidth(vvthin))
    bar(3, fcolor(dkorange) lcolor(dkorange) lwidth(vvthin))
    yscale(range(-20(20)100))
    ylabel(-20(20)100, nogrid labsize(small))
    ylabel(-20 "20" 0 "0" 20 "20" 40 "40" 60 "60" 80 "80" 100 "100")
    yline(0, lcolor(black) lwidth(vvthin))
```

5. RATES OF COLLEGE ENROLLMENT AMONG GRADUATES HIGHLY QUALIFIED TO ATTEND FOUR-YEAR COLLEGES, BY COLLEGE TYPE

```
title("Rates of Highly Qualified Students Attending College, by Race", size(medlarge) span)
    subtitle("Among Graduates Eligible to Attend Four-Year Universities", size(*.8) span)
    ytitle("Percent of Highly-Qualified Graduates" " ", size(small))
legend(region(lcolor(white)) position(12) row(1) label(1 "Enrolled at 4-Yr College")
    label(2 "Enrolled at 2-Yr College") label(3 "Not Enrolled in College") symxsize(2)
symysize(2) size(*.7))
    graphregion(color(white) fcolor(white) lcolor(white))
    plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} first-time ninth graders. Students who
transferred into or out of ${agency_name} are excluded"
"from the sample. Eligibility to attend a public four-year university is based on
students' cumulative GPA and ACT/SAT scores."
"Sample includes `count_1' African American, `count_2' Asian American students, `count_3'
Hispanic, and `count_4' White students."
"Post-secondary enrollment data are from NSC matched records. $admin_nsc_note", size(2));
#delimit cr
graph export "D5_Col_Enrl_HiQualified_by_Type.emf", replace
graph save "D5_Col_Enrl_HiQualified_by_Type.gph", replace
}
```


6. GAPS IN RATES OF COLLEGE ENROLLMENT BETWEEN LATINO HIGH SCHOOL GRADUATES AND WHITE HIGH SCHOOL GRADUATES



Purpose: This Strategic Performance Indicator explores gaps in college enrollment rates by ethnicity, before and after accounting for differences in prior academic achievement, socioeconomic status, and both of these background characteristics. While the analysis evaluates separately the college enrollment gaps between Black and White students and between Latino and White students, it can be modified to focus on the gap between any two races or ethnicities.

Required Analysis File Variables:

sid
chrt_grad
last_hs_code
race_ethnicity
test_math_8
frpl_ever
enrl_1oct_grad_yr1_any

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation.
- Keep only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).
- Drop race/ethnic groups with less than 20 students eligible to attend a four-year university. You may further restrict the sample to only include students from the most representative racial/ethnic sub-groups in your agency.

Ask Yourself

- How do racial gaps in college enrollment change after prior achievement is accounted for? How do these gaps change after socioeconomic status is accounted for?
- Do these gaps still exist when you account for both prior achievement and socioeconomic status? Do they reverse direction, suggesting that minority students enroll in college at higher rates when compared with White students with similar background characteristics?
- Do you observe differences in the degree to which the White-Black gap and the White-Latino gap decline after accounting for prior achievement and socioeconomic status? If the adjusted gap between White and Latino students, for example, is still sizeable, what additional barriers may be impeding access to college for Latino students?

6. GAPS IN RATES OF COLLEGE ENROLLMENT BETWEEN LATINO HIGH SCHOOL GRADUATES AND WHITE HIGH SCHOOL GRADUATES

Analytic Technique: Calculate the difference between the proportion of Black (or Latino) high school graduates and the proportion of White high school graduates who enrolled in college—in raw terms and after accounting for 8th grade test scores, for eligibility for Free or Reduced Price Lunch (FRPL), and for both of these characteristics.

```
/**** C. College Enrollment ****/
/**** 6. Gaps in Rates of College Enrollment Between Race/Ethnic Groups ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND have non-missing eighth grade test scores AND non-missing FRPL status

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
keep if frpl_ever != . & test_math_8 != .
```

// Step 3: Include only black, Latino, and white students

```
keep if race_ethnicity==1 | race_ethnicity == 3 | race_ethnicity == 5
gen afam = (race_ethnicity == 1)
gen hisp = (race_ethnicity == 3)
gen white = (race_ethnicity == 5)
```

// Step 4: Estimate the unadjusted and adjusted differences in college enrollment between Latino and white students and between black and white students

// 1. Create a unique codeentifier for each cohort at each high school, so that we can cluster the standard errors at the cohort/high school level

```
egen cluster_var = concat(chrt_grad last_hs_code)
```

// 2. Fit 4 separate regression models with and without control variables, and save the coefficients associated with each race.

// 2A. Estimate unadjusted enrollment gap

```
reg enrl_1oct_grad_yr1_any afam hisp, robust cluster(cluster_var)
gen afam_unadj = _b[afam]
gen hisp_unadj = _b[hisp]
```

// 2B. Estimate enrollment gap adjusting for prior achievement

```
reg enrl_1oct_grad_yr1_any afam hisp test_math_8, robust cluster(cluster_var)
gen afam_adj_prior_ach = _b[afam]
gen hisp_adj_prior_ach = _b[hisp]
```

// 2C. Estimate enrollment gap adjusting for FRPL status

```
reg enrl_1oct_grad_yr1_any afam hisp frpl_ever, robust cluster(cluster_var)
gen afam_adj_frpl = _b[afam]
gen hisp_adj_frpl = _b[hisp]
```


6. GAPS IN RATES OF COLLEGE ENROLLMENT BETWEEN LATINO HIGH SCHOOL GRADUATES AND WHITE HIGH SCHOOL GRADUATES

// 2D. Estimate enrollment gap adjusting for prior achievement and FRPL status

```
reg enrl_1oct_grad_yr1_any afam hisp frpl_ever test_math_8, robust cluster(cluster_var)
gen afam_adj_prior_frpl = _b[afam]
gen hisp_adj_prior_frpl = _b[hisp]
```

//3. Transform the regression coefficients estimated in Step 4.2 to be displayed in positive % terms

```
foreach race in afam hisp {
    replace `race'_unadj = (0 - `race'_unadj) * 100
    replace `race'_adj_prior_ach = (0 - `race'_adj_prior_ach) * 100
    replace `race'_adj_frpl = (0 - `race'_adj_frpl) * 100
    replace `race'_adj_prior_frpl = (0 - `race'_adj_prior_frpl) * 100
}
```

// Step 5: Retain a data file containing only the regression coefficients

```
keep afam_* hisp_*
duplicates drop
```

// Step 6: Prepare to graph the results

// Generate a cohort label to be used in the footnote for the graph

```
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

// Step 7: Graph the results

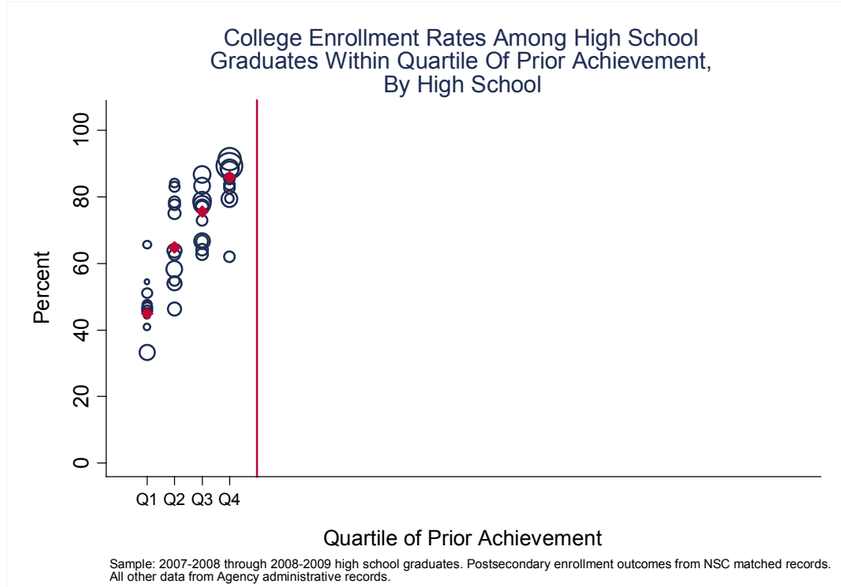
// 1. Graph results for black and white students

```
#delimit ;
graph bar afam_unadj afam_adj_prior_ach afam_adj_frpl afam_adj_prior_frpl,
    legend(row(2) size(vsmall) region(lcolor(white))
    label(1 "Unadjusted enrollment gap")
    label(2 "Gap adjusted for prior achievement")
    label(3 "Gap adjusted for FRPL status")
    label(4 "Gap adjusted for prior achievement & FRPL status"))
outergap(300)
blabel(bar, format(%9.0f) size(vsmall))
    bar(1, fcolor(dknavy) lcolor(dknavy) fi(inten100))
    bar(2, fcolor(dknavy) lcolor(dknavy) fi(inten70))
    bar(3, fcolor(dknavy) lcolor(dknavy) fi(inten50))
    bar(4, fcolor(dknavy) lcolor(dknavy) fi(inten20))
    title("Differences In Rates Of College Enrollment"
```

6. GAPS IN RATES OF COLLEGE ENROLLMENT BETWEEN LATINO HIGH SCHOOL GRADUATES AND WHITE HIGH SCHOOL GRADUATES

```
    "Between Black High School Graduates And White High"
    "School Graduates", size(med))
    ytitle("Percentage Points", margin(2 2 0 0) size(small))
    yscale(range(-20(10)50)) ylabel(-20(10)50, labsize(small))
    graphregion(color(white) fcolor(white) lcolor(white))
    plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' high school graduates. Postsecondary enrollment outcomes
from NSC matched records. All other data from ${agency_name} administrative records.",
size(vsmall));
#delimit cr

// 2. Graph results for Latino and white students
#delimit ;
graph bar hisp_unadj hisp_adj_prior_ach hisp_adj_frpl hisp_adj_prior_frpl,
    legend(row(2) size(vsmall) region(lcolor(white))
    label(1 "Unadjusted enrollment gap")
    label(2 "Gap adjusted for prior achievement")
    label(3 "Gap adjusted for FRPL status")
    label(4 "Gap adjusted for prior achievement & FRPL status"))
    outergap(300)
blabel(bar, format(%9.0f) size(vsmall))
    bar(1, fcolor(dknavy) lcolor(dknavy) fi(inten100))
    bar(2, fcolor(dknavy) lcolor(dknavy) fi(inten70))
    bar(3, fcolor(dknavy) lcolor(dknavy) fi(inten50))
    bar(4, fcolor(dknavy) lcolor(dknavy) fi(inten20))
    title("Differences In Rates Of College Enrollment"
    "Between Latino High School Graduates And White High"
    "School Graduates", size(med))
    ytitle("Percentage Points", margin(2 2 0 0) size(small))
    yscale(range(-20(10)50)) ylabel(-20(10)50, labsize(small))
    graphregion(color(white) fcolor(white) lcolor(white))
    plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' high school graduates. Postsecondary enrollment outcomes from
NSC matched records." "All other data from ${agency_name} administrative records.",
size(vsmall));
#delimit cr
graph export "D6_Col_Enrl_Gap_Latino_Black.emf", replace
graph save "D6_Col_Enrl_Gap_Latino_Black.gph", replace
}
```


7. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES - BUBBLES



Purpose: This SPI highlights the variation in college-going rates across high schools when students with similar prior achievement are compared. To conduct these comparisons, we first sort all incoming ninth-graders into quartiles based on their 8th grade test scores. We then examine college-going rates by high school among graduates within each of these quartiles.

Required Analysis File Variables:

sid
chrt_grad
last_hs_name
hs_diploma
qrt_8_math_std
enrl_1oct_grad_yr1_any

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND have non-missing eighth grade test scores.
- Drop high schools with less than 20 students in each quartile enrolled in ninth grade across the cohorts.
- Keep only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).

Ask Yourself

- How do college enrollment rates vary across high schools for students within the same quartile of 8th grade test scores (that is, when we compare students with similar prior achievement)?
- What is the difference between the high schools with the lowest and with the highest rates in each quartile?
- Are across-school differences in colleges enrollment rates particularly large for students of certain achievement profile—for example, for students with 8th grade test scores in the bottom quartile?

7. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES - BUBBLES

Analytic Technique: Calculate the share of students in each 8th grade test score quartile at each high school who enroll in college seamlessly after high school graduation.

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND have non-missing eighth grade test scores

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
keep if qrt_8_math != .
```

// Step 3: Create agency- and school-level average outcomes for each quartile

```
// 1. Calculate the mean of each outcome variable by high school
collapse (sum) enrl_1oct_grad_yr1_any hs_diploma, by(last_hs_name qrt_8_math)
gen pct_enrl = enrl_1oct_grad_yr1_any / hs_diploma * 100
// 2. Calculate the mean of each outcome variable for the agency as a whole
egen num = sum(enrl_1oct_grad_yr1_any), by(qrt_8_math)
egen denom = sum(hs_diploma), by(qrt_8_math)
gen agency_avg = num / denom * 100
drop num denom
```

// Step 4: Create a variable to identify the test score quartile

```
gen agency_quartile_code = .
forvalues qrt = 1(1)4 {
    local qrt_plot = `qrt' * 2
    replace agency_quartile_code = 1.`qrt_plot' if qrt_8_math == `qrt'
}
```

// Step 5: Prepare to graph the results

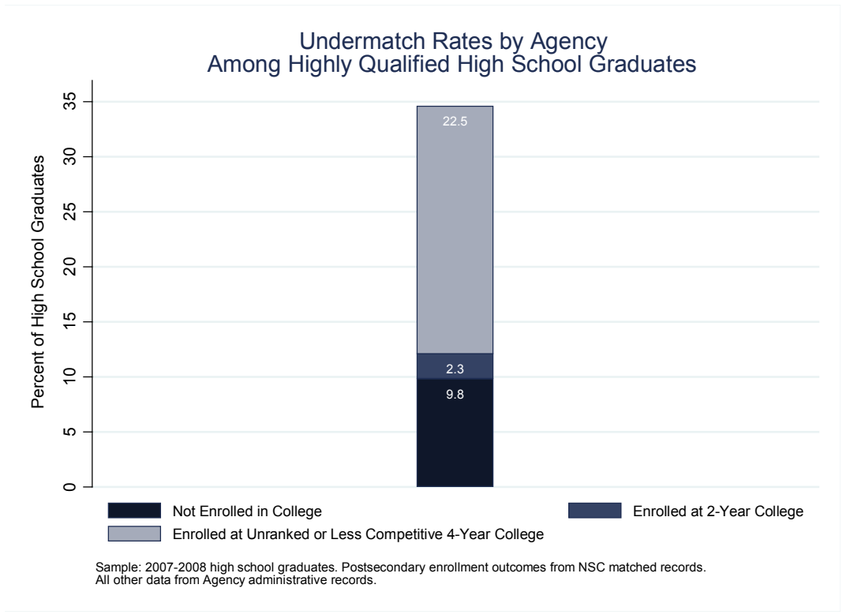
```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```


7. COLLEGE ENROLLMENT RATES BY 8TH GRADE ACHIEVEMENT QUARTILES - BUBBLES

// Step 6: Graph the results

```
#delimit ;
graph twoway scatter pct_enrl agency_quartile_code [aweight = hs_diploma],
    msymbol(Oh) msize(vsmall) mcolor(dknavy) ||
scatter agency_avg agency_quartile_code,
    mcolor(cranberry) msymbol(D) msize(small)
title("College Enrollment Rates Among High School"
"Graduates Within Quartile Of Prior Achievement,"
"By High School", size(med))
    xscale(range(1 6)) yscale(range(0 105)) ylabel(0 20 40 60 80 100)
    xlabel(1.2 "Q1" 1.4 "Q2" 1.6 "Q3" 1.8 "Q4", labsize(small))
    xtitle(" " "Quartile of Prior Achievement") ytitle("Percent" " ")
    ylabel(,nogrid) legend(off)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
xline(2)
note("Sample: `chrt_label' high school graduates. Postsecondary enrollment outcomes from
NSC matched records."
"All other data from ${agency_name} administrative records.", size(vsmall));
#delimit cr
graph export "D7_Col_Enrl_by_Eighth_Qrt_Bubbles.emf", replace
graph save "D7_Col_Enrl_by_Eighth_Qrt_Bubbles.gph", replace
}
```

8. UNDERMATCH RATES AMONG HIGHLY QUALIFIED HIGH SCHOOL GRADUATES



Purpose: This Strategic Performance Indicator examines the prevalence of "undermatch" in the agency—that is, the extent to which high school graduates with strong academic records pursue enrollment in colleges and universities less selective than those for which they are likely qualified. The SPI does so by illustrating the rates at which highly qualified graduates are enrolling at 2-year colleges, less competitive 4-year colleges, or forgoing college altogether, instead of pursuing selective colleges that may provide a better academic and social fit for these students' potential, ambition, and preparation.

Required Analysis File Variables:

sid
chrt_grad
highly_qualified
first_college_opeid_4yr
enrl_1oct_grad_yr1_4yr
enrl_1oct_grad_yr1_2yr
enrl_1oct_grad_yr1_any

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation.
- Keep only highly qualified high school graduates (i.e. students who have obtained a high school diploma on time with 1) a cumulative GPA of 3.0 or higher and Math/Verbal SAT score of 1300 or higher, or 2) a cumulative GPA of 3.3 or higher and Math/Verbal SAT score of 1200 or higher, or 3) a cumulative GPA of 3.7 or higher and Math/Verbal SAT score of at least 1100).
- Keep only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).

8. UNDERMATCH RATES AMONG HIGHLY QUALIFIED HIGH SCHOOL GRADUATES

A Note on College Selectivity

To determine the selectivity of the postsecondary institutions in which high school graduates enroll, we typically rely on Barron's College Rankings. Barron's has developed well-established college selectivity ratings based on the degree of admissions competitiveness at four-year colleges and universities. Factors used in determining these rankings include the median SAT and ACT scores, high school class rankings, and grade point average among incoming college freshmen. The seven selectivity rankings Barron's assigns are "Most Competitive," "Highly Competitive," "Very Competitive," "Competitive," "Less Competitive," "Non-Competitive," and "Special."

As part of this exercise, we have provided a simplified table from which the selectivity ratings of the colleges and universities included in this dataset can be obtained. In conducting this analysis for your own agency, you need to select a source of college selectivity ratings, such as Barron's, and use it in place of the college selectivity table used in this exercise.

Ask Yourself

- How do college enrollment rates vary across high schools for students within the same quartile of 8th grade test scores (that is, when we compare students with similar prior achievement)?
- What is the difference between the high schools with the lowest and with the highest rates in each quartile?
- Are across-school differences in colleges enrollment rates particularly large for students of certain achievement profile—for example, for students with 8th grade test scores in the bottom quartile?

8. UNDERMATCH RATES AMONG HIGHLY QUALIFIED HIGH SCHOOL GRADUATES

Analytic Technique: Calculate the proportion of highly qualified graduates who do not enroll in college, enroll in 2-year college, and enroll in least competitive and unranked 4-year colleges the fall following high school graduation.

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation AND are highy qualified

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_begin}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
keep if highly_qualified == 1
```

// Step 3: Link the analysis file with the college selectivity table to obtain the selectivity level for each college. Use this selectivity information to create college enrollment indicator variables for each college selectivity level. This script assumes that there are 5 levels of selectivity, as in Barron's College Rankings—Most Competitive (1), Highly Competitive (2), Very Competitive (3), Competitive (4), Least Competitive (5)—as well as a category for colleges without assigned selectivity (assumed to be not competitive).

```
// 1. Link analysis file with college selectivity data
rename first_college_opeid_4yr college_id
merge m:1 college_id using college_selectivity.dta, keep(1 3) keepusing(rank) nogen

// 2. Create college enrollment dummy variables for each of the five selectivity levels
forvalues i = 1/5 {
    gen enrl_1oct_grad_yr1_4yr_`i' = (enrl_1oct_grad_yr1_4yr == 1 & rank == `i')
}
// 3. Create a college enrollment dummy variable for colleges that are not ranked
gen enrl_1oct_grad_4yr_nr = (enrl_1oct_grad_yr1_4yr == 1 & (rank == 6 | rank ==. ))
// 4. Rename and label the college enrollment variables with clear labels
rename enrl_1oct_grad_yr1_4yr_1 enrl_1oct_grad_4yr_mc
rename enrl_1oct_grad_yr1_4yr_2 enrl_1oct_grad_4yr_hc
rename enrl_1oct_grad_yr1_4yr_3 enrl_1oct_grad_4yr_vc
rename enrl_1oct_grad_yr1_4yr_4 enrl_1oct_grad_4yr_c
rename enrl_1oct_grad_yr1_4yr_5 enrl_1oct_grad_4yr_lc

label var enrl_1oct_grad_4yr_mc "Enrolled at Most Competitive College Fall After HS Grad"
label var enrl_1oct_grad_4yr_hc "Enrolled at Highly Competitive College Fall After HS
Grad"
label var enrl_1oct_grad_4yr_vc "Enrolled at Very Competitive College Fall After HS Grad"
label var enrl_1oct_grad_4yr_c "Enrolled at Competitive College Fall After HS Grad"
label var enrl_1oct_grad_4yr_lc "Enrolled at Least Competitive College Fall After HS Grad"
label var enrl_1oct_grad_4yr_nr "Enrolled at Non-Competitive College Fall After HS Grad"
//5. Check to make sure that each student who appears enrolled in college as of the first fall after high school graduation is associated with one and only one college selectivity level
assert enrl_1oct_grad_4yr_mc + enrl_1oct_grad_4yr_hc + enrl_1oct_grad_4yr_vc + enrl_1oct_
grad_4yr_c + enrl_1oct_grad_4yr_lc + enrl_1oct_grad_4yr_nr == 1 if enrl_1oct_grad_yr1_4yr
== 1
```


8. UNDERMATCH RATES AMONG HIGHLY QUALIFIED HIGH SCHOOL GRADUATES

// Step 4: Create undermatch outcomes

```
//1. Not enrolled in college
gen no_college = (enrl_1oct_grad_yr1_any == 0)
//2. Enrolled in a 2-year college
gen enrl_2yr = (enrl_1oct_grad_yr1_2yr == 1)
//3. Enrolled in a least competitive 4-year college or a 4-year college without an assigned selectivity
gen enrl_4yr_under = (enrl_1oct_grad_4yr_nr == 1)
replace enrl_4yr_under = 1 if enrl_1oct_grad_4yr_lc == 1
//4. Enrolled in a 4-year college with a selectivity rating of Competitive, Very Competitive, Most Competitive, or Highly Competitive
gen enrl_4yr_match = (enrl_1oct_grad_4yr_c == 1 | enrl_1oct_grad_4yr_vc == 1 | enrl_1oct_
grad_4yr_hc == 1 | enrl_1oct_grad_4yr_mc == 1)
//5. Check to make sure that each student is associated one and only one undermatch outcome
// assert no_college + enrl_2yr + enrl_4yr_under + enrl_4yr_match == 1
```

// Step 5: Create agency-average undermatch outcomes and transform them into % terms

```
collapse (mean) no_college enrl_2yr enrl_4yr_under enrl_4yr_match (count) N = sid
foreach v of varlist no_college enrl_2yr enrl_4yr_under enrl_4yr_match {
    replace `v' = round(`v' * 100, 0.1)
}
```

// Step 6: Prepare to graph the results

```
// Generate a cohort label to be used in the footnote for the graph
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

// Step 7: Graph the results

```
#delimit ;
graph bar no_college enrl_2yr enrl_4yr_under, stack
    blabel(bar, format(%9.1f) size(2.05) position(inside) color(white))
    bar(1, fcolor(dknavy) lcolor(dknavy) finten(200) lwidth(thin))
    bar(2, fcolor(dknavy) lcolor(dknavy) finten(90) lwidth(thin))
    bar(3, fcolor(dknavy) lcolor(dknavy) finten(40) lwidth(thin))
    yscale(range(0(5)35)) outergap(400)
    ylabel(0(5)35, labsize(small))
title("Undermatch Rates by Agency"
    "Among Highly Qualified High School Graduates", size(med))
    ytitle("Percent of High School Graduates" " ", size(small))
```

D. College Enrollment

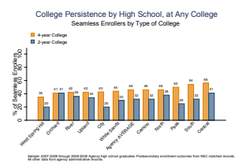
8. UNDERMATCH RATES AMONG HIGHLY QUALIFIED HIGH SCHOOL GRADUATES

```
legend(region(lcolor(white))
    label(1 "Not Enrolled in College")
    label(2 "Enrolled at 2-Year College")
    label(3 "Enrolled at Unranked or Less Competitive 4-Year College")
    symxsize(*.7) symysize(*.7) size(*.7))
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note("Sample: `chrt_label' high school graduates. Postsecondary enrollment outcomes from
NSC matched records."
"All other data from ${agency_name} administrative records.", size(vsmall)) ;
#delimit cr
graph export "D8_Undermatching_HiQualified.emf", replace
graph save "D8_Undermatching_HiQualified.gph", replace
}
```


E. COLLEGE PERSISTENCE

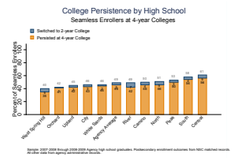
For many high school graduates, college enrollment is just the first of many hurdles on the road to postsecondary success. While considerable attention has been paid to challenges that surround college preparedness, access, and enrollment, only recently has conversation expanded to consider barriers to degree completion. These barriers must be understood and addressed at both the secondary and postsecondary levels for college attainment rates to increase. In the last section of the education pipeline, you examine patterns of persistence to the second year of college to identify early indications of student progress towards degree attainment.

To explore college persistence, use the models below:



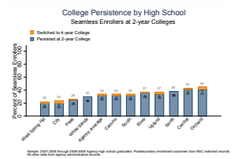
1. PERSISTENCE RATES TO THE SECOND YEAR OF COLLEGE BY HIGH SCHOOL

Provides an agency snapshot of persistence to the second year of college by examining persistence rates across high schools in the system. The analysis also illuminates differences in persistence by level of college first attended (two-year vs. four-year) and by time of initial entry (seamless vs. delayed enrollment).



2. PERSISTENCE ACROSS TWO-YEAR AND FOUR-YEAR COLLEGES

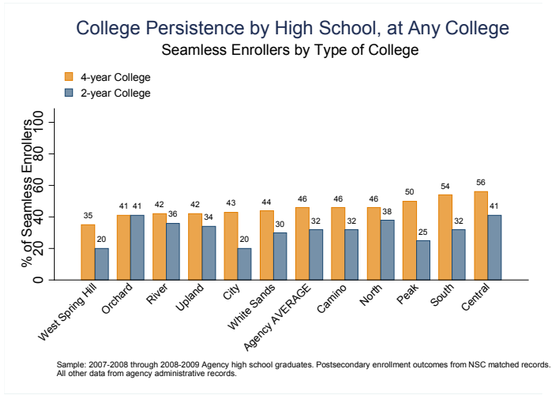
Provides a snapshot of persistence to the second year of college from one type of college to another for different high schools in the system

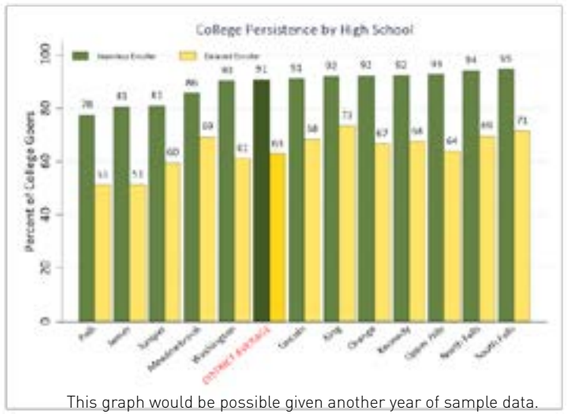


3. TOP-ENROLLING COLLEGES/UNIVERSITIES OF AGENCY GRADUATES

Reports enrollment and persistence rates among top-enrolling two- and four-year higher education institutions attended by agency graduates.

1. PERSISTENCE RATES TO THE SECOND YEAR OF COLLEGE BY HIGH SCHOOL





This graph would be possible given another year of sample data.

Purpose: Initial enrollment decisions can dramatically affect higher education trajectories and the likelihood of degree attainment. This analysis provides a snapshot of persistence to the second year of college by examining persistence rates across high schools in the system. The analysis illuminates differences in persistence by level of college first attended (two-year vs. four-year). Given another year of sample data, the analysis could also be conducted by time of initial entry (seamless vs. delayed enrollment).

Required Analysis File Variables:

sid
enrl_1oct_grad_yr1_any
enrl_1oct_grad_yr1_2yr
enrl_1oct_grad_yr1_4yr
enrl_grad_persist_any
enrl_grad_persist_2yr
enrl_grad_persist_4yr
last_hs_code
last_hs_name
enrl_ever_w2_grad_any (given another year of sample data)

Analysis-Specific Sample Restrictions:

- Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation
- Keep only graduates who received regular or advanced diplomas (i.e. exclude students who received SPED diplomas and other certificates).
- Drop high schools with less than 20 students in the sample.

Ask Yourself

- How does college persistence for enrollers at 2-year colleges compare to enrollers at 4-year colleges? Given another year of sample data, how does college persistence for seamless enrollers compare to delayed enrollers?

Possible Next Steps or Action Plans: Consider establishing MOUs with local community colleges to obtain detailed data on graduates' postsecondary pursuits at two-year colleges (Course enrollment and transcript data) allowing agencies to explore persistence rates by assignment to remediation coursework.

1. PERSISTENCE RATES TO THE SECOND YEAR OF COLLEGE BY HIGH SCHOOL

Analytic Technique: Calculate the proportion of students who persist to the second year of college by the high school those students first attended.

```
/**** E. College Persistence ****/
/**** 1. Persistence Rates to the Second Year of College by High School ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
```

// Step 3: Rename outcome variable names for simplicity

```
rename enrl_grad_persist_any persist_any
rename enrl_grad_persist_2yr persist_2yr
rename enrl_grad_persist_4yr persist_4yr
rename enrl_1oct_grad_yr1_any enrl_any
rename enrl_1oct_grad_yr1_2yr enrl_2yr
rename enrl_1oct_grad_yr1_4yr enrl_4yr
```

// Step 4: Obtain the agency-level average for persistence and enrollment

```
preserve
    collapse (sum) persist_any enrl_4yr if enrl_4yr==1
    tempfile agency_level_4yr
    save `agency_level_4yr'
restore

preserve
    collapse (sum) persist_any enrl_2yr if enrl_2yr==1
    tempfile agency_level_2yr
    save `agency_level_2yr'
restore
```

// Step 5: Obtain the school-level average for persistence and enrollment

```
preserve
    collapse (sum) persist_any enrl_4yr if enrl_4yr==1, by(last_hs_code last_hs_name)
    tempfile school_level_4yr
    save `school_level_4yr'
restore

collapse (sum) persist_any enrl_2yr if enrl_2yr==1, by(last_hs_code last_hs_name)
```

// Step 6: Append on the previous agency-level and school-level files

```
append using `agency_level_4yr'
append using `agency_level_2yr'
append using `school_level_4yr'
```

1. PERSISTENCE RATES TO THE SECOND YEAR OF COLLEGE BY HIGH SCHOOL

// Step 7: Provide a hs name label for the agency average and shorten hs name

```
replace last_hs_code = 0 if last_hs_code == .
replace last_hs_name = "${agency_name} AVERAGE" if mi(last_hs_name)
replace last_hs_name = subinstr(last_hs_name, " High School", "", .)
```

// Step 8: Calculate percent persistence at 4-year and 2-year colleges and multiply outcomes of interest by 100 for graphical representation of the rates

```
gen pct_persist_any = persist_any / enrl_4yr
replace pct_persist_any  = round((pct_persist_any  * 100))

foreach type in 2yr 4yr {
    gen pct_persist_`type' = persist_any / enrl_`type'
    replace pct_persist_`type' = round((pct_persist_`type' * 100))
}
```

// Step 9: Drop any high schools with fewer than 20 students

```
drop if enrl_4yr < 20
drop if enrl_2yr < 20
```

// Step 10: Consolidate persistence data into single column and then reshape the data

```
gen n=2 if  pct_persist_2yr~=.
replace n=4 if  pct_persist_4yr~=.
replace pct_persist_4yr= pct_persist_2yr if  pct_persist_4yr==.
keep last_hs_code last_hs_name pct_persist_4yr n
rename pct_persist_4yr pct_persist

reshape wide pct_persist, i(last_hs_name) j(n)
```

// Step 11: Prepare to graph the results

// Generate a cohort label to be used in the footnote for the graph

```
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

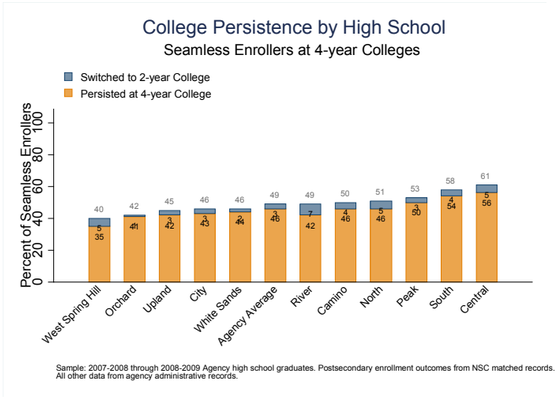
// Step 12: Graph the results

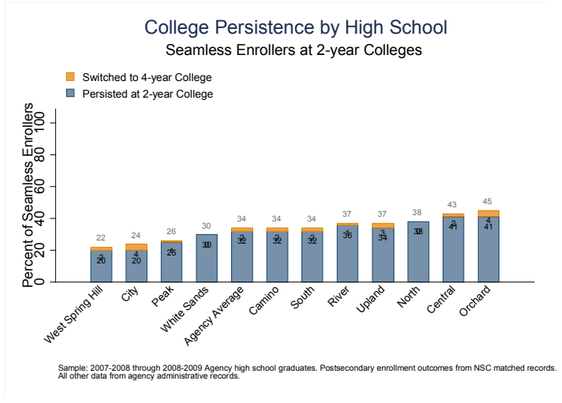
```
#delimit ;
graph bar pct_persist4 pct_persist2,
    over(last_hs_name, label(angle(45)labsize(small)) sort(pct_persist4)) bargap(0)
outergap(100)
    bar(1, fcolor(dkorange) fi(inten70) lcolor(dkorange) lwidth(vvvthin))
    bar(2, fcolor(navy) fi(inten60) lcolor(navy) lwidth(vvvthin))
```


1. PERSISTENCE RATES TO THE SECOND YEAR OF COLLEGE BY HIGH SCHOOL

```
    blabel(total, position(outside) color(black) size(vsmall) format(%8.0f))
legend(label(1 "4-year College") label(2 "2-year College")
    position(11) ring(1) symxsize(2) symysize(2) rows(2) size(small) region(lstyle(none)
lcolor(none) color(none)))
title("College Persistence by High School, at Any College")
    subtitle("Seamless Enrollers by Type of College")
    ytitle("% of Seamless Enrollers")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} high school graduates. Postsecondary
enrollment outcomes from NSC matched records."
"All other data from agency administrative records.", size(vsmall));
#delimit cr
graph export "E1_Persistence_by_HS.emf", replace
graph save "E1_Persistence_by_HS.gph", replace
}
```

2. PERSISTENCE ACROSS TWO-YEAR AND FOUR-YEAR COLLEGES





Purpose: This analysis provides a snapshot of persistence to the second year of college from one type of college to another for different high schools in the system. The left analysis charts explores how seamless enrollers in 4-year colleges either persist at a 4-year or switch to a 2-year. The right analysis charts how seamless enrollers in 2-year colleges either persist at a 2-year or switch to a 4-year.

Required Analysis File Variables:

sid
enrl_1oct_grad_yr1_any
enrl_1oct_grad_yr1_2yr
enrl_1oct_grad_yr1_4yr
enrl_grad_persist_any
enrl_grad_persist_2yr
enrl_grad_persist_4yr
last_hs_code
last_hs_name

Analysis-Specific Sample Restrictions:

- Keep the three most recent cohorts of graduates for which persistence in college over four consecutive years can be reported.
- Keep only graduates enrolled in 4-yr colleges and universities the fall following high school graduation.
- Keep only graduates for whom cumulative high school GPAs can be calculated (or obtained from the agency)
- Only include the top six enrolling 4-year colleges
- Only report persistence rates among students falling in each high school GPA category if the sample includes 25 or more students

Ask Yourself

- How do the rates of persistence or switching differ for seamless enrollers at 4-year vs. 2-year colleges?

Possible Next Steps or Action Plans: Create individual school-level reports for administrators and college counselors to communicate which postsecondary institutions are associated with greater rates of persistence. Additionally, conduct similar analyses that include more detailed institutional information that may be associated with students' prospects of persisting (e.g. cost of tuition and room/board, financial aid, etc.).

2. PERSISTENCE ACROSS TWO-YEAR AND FOUR-YEAR COLLEGES

Analytic Technique: Calculate the proportion of 4-yr college-goers who persist through four years of college by the postsecondary institution first attended and cumulative high school GPA category.

```
/**** E. College Persistence ****/
/**** 3. Persistence Across Two-Year and Four-Year Colleges ****/
{
```

// Step 1: Load the college-going analysis file into Stata

```
use "CG_Analysis", clear
```

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation

```
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')
```

// Step 3: Rename outcome variable names for simplicity

```
rename enrl_grad_persist_2yr persist_2yr
rename enrl_1oct_grad_yr1_2yr enrl_2yr
rename enrl_grad_persist_4yr persist_4yr
rename enrl_1oct_grad_yr1_4yr enrl_4yr
```

// Step 4: Create binary outcomes for enrollers who switch from 4-yr to 2-yr, or vice versa

```
gen persist_4to2yr = (enrl_4yr == 1 & enrl_1oct_grad_yr2_2yr == 1) if !mi(chrt_grad)
gen persist_2to4yr = (enrl_2yr == 1 & enrl_1oct_grad_yr2_4yr == 1) if !mi(chrt_grad)
```

// Step 5: Obtain the agency-level average for the different persistence outcomes

```
preserve
    collapse (sum) persist_4yr persist_4to2yr enrl_4yr persist_2yr persist_2to4yr enrl_2yr
    tempfile agency_level
    save `agency_level'
restore
```

// Step 6: Obtain the school-level average for the different persistence outcomes

```
collapse (sum) persist_4yr persist_4to2yr enrl_4yr persist_2yr persist_2to4yr enrl_2yr,
by(last_hs_code last_hs_name)
append using `agency_level'
```

// Step 7: Provide a hs name label for the agency average and shorten hs name

```
replace last_hs_name = "${agency_name} Average" if mi(last_hs_name)
replace last_hs_code = 0 if mi(last_hs_code)
replace last_hs_name = subinstr(last_hs_name, " High School", "", .)
```

// Step 8: Generate percentages for different persistence outcomes. Multiply outcomes of interest by 100 for graphical representations of the rates

```
gen pct_persist_4yr = persist_4yr / enrl_4yr
gen pct_persist_4to2yr = persist_4to2yr / enrl_4yr
gen pct_persist_2yr = persist_2yr / enrl_2yr
gen pct_persist_2to4yr = persist_2to4yr / enrl_2yr
```

2. PERSISTENCE ACROSS TWO-YEAR AND FOUR-YEAR COLLEGES

```
foreach var in pct_persist_2yr pct_persist_2to4yr pct_persist_4yr pct_persist_4to2yr {
    replace `var' = round((`var' * 100))
}
```

// Step 9: Create total persistence rates by summing up the other variables

```
gen total_persist_4yr =  pct_persist_4yr + pct_persist_4to2yr
gen total_persist_2yr =  pct_persist_2yr + pct_persist_2to4yr
```

//Step 10: Prepare to graph the results

// 1. Generate a cohort label to be used in the footnote for the graph

```
local temp_begin = `chrt_grad_begin'-1
local temp_end = `chrt_grad_end'-1
if `chrt_grad_begin'==`chrt_grad_end' {
    local chrt_label "`temp_begin'-`chrt_grad_begin'"
}
else {
    local chrt_label "`temp_begin'-`chrt_grad_begin' through `temp_end'-`chrt_grad_end'"
}
```

// 2. Generate graphing code to place value labels for the total persistence rates; change xpos (the position of the first leftmost label) and xposwidth (the horizontal width of the labels) to finetune.

```
foreach yr in 4 2 {
    sort total_persist_`yr'yr
    local total_persist_`yr'yr ""
    local num_obs = _N
    foreach n of numlist 1/`num_obs' {
    local temp_total_persist_`yr'yr = total_persist_`yr'yr in `n'
    local total_persist_`yr'yr `"`total_persist_`yr'yr' `temp_total_persist_`yr'yr'"'
    }
    local total_persist_`yr'yr_label ""
    local xpos = 7
    local xposwidth = 93.5
    foreach val of local total_persist_`yr'yr {
    local val_pos = `val' + 6
    local total_persist_`yr'yr_label `"`total_persist_`yr'yr_label' text(`val_pos' `xpos'
"`val'", size(2.1) color(gs7))"'
    local xpos = `xpos' + `xposwidth'/_N
    }
    disp `"`total_persist_`yr'yr_label'"'
}
```

// Step 11: Graph the results (1/2) for seamless enrollers at 4-year colleges

```
#delimit ;
graph bar pct_persist_4yr pct_persist_4to2yr if enrl_4yr >= 20,
    over(last_hs_name, label(angle(45)labsize(small)) sort(total_persist_4yr)) bargap(0)
outergap(100)
    bar(1, fcolor(dkorange) fi(inten70) lcolor(dkorange) lwidth(vvvthin))
    bar(2, fcolor(navy) fi(inten60) lcolor(navy) lwidth(vvvthin)) stack
    blabel(bar, position(inside) color(black) size(vsmall) format(%8.0f))
```


2. PERSISTENCE ACROSS TWO-YEAR AND FOUR-YEAR COLLEGES

```
legend(label(1 "Persisted at 4-year College") label(2 "Switched to 2-year College")
    position(11) order(2 1) ring(1) symxsize(2) symysize(2) rows(2) size(small)
region(lstyle(none) lcolor(none) color(none)))
title("College Persistence to 2nd Year, by High School")
    subtitle("Seamless Enrollers at 4-year Colleges")
    `total_persist_4yr_label'
    ytitle("Percent of Seamless Enrollers")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} high school graduates. Postsecondary
enrollment outcomes from NSC matched records."
"All other data from agency administrative records.", size(vsmall));
#delimit cr
graph export "E2a_Persistence_4yr_Seamless_Enrlers.emf", replace
graph save "E2a_Persistence_4yr_Seamless_Enrlers.gph", replace
```

// Step 12: Graph the results (1/2) for seamless enrollers at 2-year colleges

```
#delimit ;
graph bar pct_persist_2yr pct_persist_2to4yr if enrl_4yr >= 20,
    over(last_hs_name, label(angle(45)labsize(small)) sort(total_persist_2yr)) bargap(0)
outergap(100)
    bar(1, fcolor(navy) fi(inten60) lcolor(navy) lwidth(vvvthin))
    bar(2, fcolor(dkorange) fi(inten70) lcolor(dkorange) lwidth(vvvthin)) stack
    blabel(bar, position(inside) color(black) size(vsmall) format(%8.0f))
legend(label(2 "Switched to 4-year College") label(1 "Persisted at 2-year College")
    position(11) order(2 1) ring(1) symxsize(2) symysize(2) rows(2) size(small)
region(lstyle(none) lcolor(none) color(none)))
title("College Persistence by High School")
    subtitle("Seamless Enrollers at 2-year Colleges")
    `total_persist_2yr_label'
    ytitle("Percent of Seamless Enrollers")
    yscale(range(0(20)100))
    ylabel(0(20)100, nogrid)
graphregion(color(white) fcolor(white) lcolor(white))
plotregion(color(white) fcolor(white) lcolor(white))
note(" " "Sample: `chrt_label' ${agency_name} high school graduates. Postsecondary
enrollment outcomes from NSC matched records."
"All other data from agency administrative records.", size(vsmall));
#delimit cr
graph export "E2b_Persistence_2yr_Seamless_Enrlers.emf", replace
graph save "E2b_Persistence_2yr_Seamless_Enrlers.gph", replace
}
```

3. TOP-ENROLLING COLLEGES/UNIVERSITIES OF AGENCY GRADUATES

Table 2

College_Name	Number_Enrolled	Percent_Enrolled	Number_ Persisted	Percent_ Persisted
UNIVERSITY OF GH	307	13	138	45
ABC STATE UNIVERSITY	273	11.5	127	46.5
UNIVERSITY XYZ	228	9.6	110	48.2
UNIVERSITY - CAMPUS 2	199	8.4	85	42.7
COLLEGE OF XYZ	145	6.1	73	50.3
OTHER 4-YEAR COLLEGES	1213	51.3	1213	
ALL 4-YEAR COLLEGES	2365	100	1088	

Table 3

College_Name	Number_Enrolled	Percent_Enrolled	Number_ Persisted	Percent_ Persisted
UVW COMMUNITY COLLEGE	115	10.5	39	33.9
COMMUNITY COLLEGE B	108	9.8	31	28.7
COMMUNITY COLLEGE 400	78	7.1	31	39.7
D COMMUNITY COLLEGE	59	5.4	18	30.5
XYZ COMMUNITY COLLEGE	58	5.3	17	29.3
OTHER 2-YEAR COLLEGES	680	61.8	680	
ALL 2-YEAR COLLEGES	1098	100	360	

Purpose: This analysis reports enrollment and persistence rates among top-enrolling two- and four-year institutions attended by graduates. This analysis illuminates differences in persistence rates to the second year of college among top-enrolling postsecondary institutions. Agency staff that advise students during their senior year may find this information useful when meeting to weigh college options.

Required Analysis File Variables:

sid
chrt_grad
enrl_1oct_grad_yr1_any
enrl_1oct_grad_yr1_2yr
enrl_1oct_grad_yr1_4yr
enrl_grad_persist_any
enrl_grad_persist_2yr
enrl_grad_persist_4yr
first_college_name_any
first_college_name_2yr
first_college_name_4yr

Analysis-Specific Sample Restrictions:

- Keep only the most recent cohort of seamless college-goers for which persistence to the second year of college can be reported
- Only include postsecondary institutions with 25 or more agency graduates attending.

Ask Yourself

- What are the top enrolling 4-year and 2-year colleges or universities in your agency? What are the persistence rates at those colleges and universities?

3. TOP-ENROLLING COLLEGES/UNIVERSITIES OF AGENCY GRADUATES

Analytic Technique: Calculate the proportion of college-goers attending top-enrolling 2- and 4-year institutions, as well as the proportion of seamless enrollers who persist to the second year of any college, by the postsecondary institution graduates first attended.

```
/**** E. College Persistence ****/
/**** 3. Top-Enrolling Colleges/Universities of Agency Graduates ****/
{
// Step 1: Load the college-going analysis file into Stata
use "${analysis}/CG_Analysis", clear

// Step 2: Keep students in high school graduation cohorts you can observe enrolling in college the fall after graduation
local chrt_grad_begin = ${chrt_grad_begin}
local chrt_grad_end = ${chrt_grad_end}
keep if (chrt_grad >= `chrt_grad_begin' & chrt_grad <= `chrt_grad_end')

// Step 3: Indicate the number of top-enrolling institutions you would like listed
local num_inst = 5

// Step 4: Calculate the number and % of students enrolled in each college the fall after graduation, and the number and % of students persisting, by college type

// 1. Calculate for 4-year colleges
preserve

      collapse (sum) enrl_1oct_grad_yr1_4yr enrl_grad_persist_4yr, by(first_college_
name_4yr)
      keep if !mi(first_college_name_4yr)

      egen total_enrolled = sum(enrl_1oct_grad_yr1_4yr)
      egen total_persisted = sum(enrl_grad_persist_4yr)

      gen pct_enrolled_college = round((enrl_1oct_grad_yr1_4yr / total_enrolled)*100, .1)
      gen pct_persist_college = round((enrl_grad_persist_4yr / enrl_1oct_grad_
yr1_4yr)*100, .1)
      format pct* %3.1f

      rename enrl_1oct_grad_yr1_4yr enrl_1oct_grad_yr1
      rename enrl_grad_persist_4yr enrl_grad_persist

      rename first_college_name_4yr first_college_name

      gen type = "4yr"

      tempfile college_4yr
      save `college_4yr'

restore
```

3. TOP-ENROLLING COLLEGES/UNIVERSITIES OF AGENCY GRADUATES

```
// 2. Calculate for 2-year colleges, and append the information for 4-year colleges
collapse (sum) enrl_1oct_grad_yr1_2yr enrl_grad_persist_2yr, by(first_college_name_2yr)
keep if !mi(first_college_name_2yr)

egen total_enrolled = sum(enrl_1oct_grad_yr1_2yr)
egen total_persisted = sum(enrl_grad_persist_2yr)

gen pct_enrolled_college = round((enrl_1oct_grad_yr1_2yr / total_enrolled)*100, .1)
gen pct_persist_college = round((enrl_grad_persist_2yr / enrl_1oct_grad_yr1_2yr)*100, .1)
format pct* %3.1f

rename enrl_1oct_grad_yr1_2yr enrl_1oct_grad_yr1
rename enrl_grad_persist_2yr enrl_grad_persist
rename first_college_name_2yr first_college_name

gen type = "2yr"

append using `college_4yr'
```

// Step 5: Create Table 1 with all 2-year and 4-year colleges listed

```
preserve
      // 1. Create two observations, one for each college type
      local newrows = _N+2
      set obs `newrows'
      replace  type="2yr" if _n==_N-1
      replace  type="4yr" if _n==_N

      replace  first_college_name = "ALL 2-YEAR COLLEGES" if type=="2yr" & mi(first_college_
name)
      replace  first_college_name = "ALL 4-YEAR COLLEGES" if type=="4yr" & mi(first_college_
name)

      // 2. Populate the new observations
      foreach type in 2 4 {
            summ total_enrolled if type == "`type'yr"
            replace enrl_1oct_grad_yr1 = r(mean) if first_college_name=="ALL `type'-YEAR
COLLEGES"
            summ total_persisted if type == "`type'yr"
            replace enrl_grad_persist = r(mean) if first_college_name=="ALL `type'-YEAR
COLLEGES"
      }
      replace pct_enrolled_college = 100 if mi(pct_enrolled_college)

      // 3. Retain, reorder, and rename necessary variables
```


3. TOP-ENROLLING COLLEGES/UNIVERSITIES OF AGENCY GRADUATES

```
      keep first_college_name enrl_1oct_grad_yr1 enrl_grad_persist pct_enrolled_college
pct_persist_college type
      order first_college_name enrl_1oct_grad_yr1 pct_enrolled_college enrl_grad_persist
pct_persist_college type

      gen rank = (regexm(first_college_name, "ALL"))
      gsort -type rank -enrl_1oct_grad_yr1
      drop rank type

      rename first_college_name College_Name
      rename enrl_1oct_grad_yr1 Number_Enrolled
      rename pct_enrolled_college Percent_Enrolled
      rename enrl_grad_persist Number_Persisted
      rename pct_persist_college Percent_Persisted

      // 4. Outsheet Table 1 into a csv file
      outsheet using "${graphs}/E3_Top_Enrl_Col_Institutions_Table_1.csv", comma replace

restore
```

// Step 6: Create Tables 2 and 3 with the number of institutions you wanted to list in Step 3 for 4-year (Table 2) and 2-year (Table 3) colleges, respectively

```
// 1. Identify the five top-enrolling 2- and 4-year institutions (5 based on the number you selected in step 3)
gsort type -pct_enrolled_college
gen rank_2yr = _n in 1/`num_inst'

gsort -type -pct_enrolled_college
gen rank_4yr = _n in 1/`num_inst'

// 2. Calculate the remaining proportion of students attending other 2- and 4-year colleges for purposes of populating the "Other" line (all other 2- and 4-year colleges beyond the number selected) in the table.
foreach type in 2yr 4yr {
      egen other_number_`type'_temp = sum(enrl_1oct_grad_yr1) if mi(rank_`type') &
type=="`type'"
      egen other_number_`type' = max(other_number_`type'_temp)
      egen other_pct_`type'_temp = sum(pct_enrolled_college) if mi(rank_`type') &
type=="`type'"
      egen other_pct_`type' = max( other_pct_`type'_temp)
      egen other_number_persist_`type'_temp = sum(enrl_grad_persist) if mi(rank_`type') &
type=="`type'"
      egen other_number_persist_`type' = max(other_number_`type'_temp)
      drop *_temp
}

keep if !mi(rank_2yr) | !mi(rank_4yr)

// 3. Create four new rows, one per college type for total counts and one per college type for colleges other than the top-enrolling ones
local newrows = _N+4
```

3. TOP-ENROLLING COLLEGES/UNIVERSITIES OF AGENCY GRADUATES

```
set obs `newrows'
replace  type="2yr" if _n==_N-2 | _n==_N-3
replace  type="4yr" if _n==_N | _n==_N-1

replace first_college_name = "OTHER 2-YEAR COLLEGES" if type=="2yr" & mi(first_college_name)
& _n==_N-3
replace first_college_name = "ALL 2-YEAR COLLEGES" if type=="2yr" & mi(first_college_name)
& _n==_N-2
replace first_college_name = "OTHER 4-YEAR COLLEGES" if type=="4yr" & mi(first_college_name)
& _n==_N-1
replace first_college_name = "ALL 4-YEAR COLLEGES" if type=="4yr" & mi(first_college_name)
& _n==_N

// 4. Populate the new rows
foreach type in 2 4 {
      summ total_enrolled if type == "`type'yr"
      replace enrl_1oct_grad_yr1 = r(mean) if first_college_name=="ALL `type'-YEAR
COLLEGES"
      summ total_persisted if type == "`type'yr"
      replace enrl_grad_persist = r(mean) if first_college_name=="ALL `type'-YEAR COLLEGES"
      summ other_number_`type'yr if type == "`type'yr"
      replace enrl_1oct_grad_yr1 = r(mean) if first_college_name=="OTHER `type'-YEAR
COLLEGES"
      summ other_pct_`type'yr if type == "`type'yr"
      replace pct_enrolled_college = r(mean) if first_college_name=="OTHER `type'-YEAR
COLLEGES"
      summ other_number_persist_`type'yr if type == "`type'yr"
      replace enrl_grad_persist = r(mean) if first_college_name=="OTHER `type'-YEAR
COLLEGES"
}
replace pct_enrolled_college = 100 if mi(pct_enrolled_college)

// 5. Retain, reorder, and rename necessary variables
keep first_college_name enrl_1oct_grad_yr1 enrl_grad_persist pct_enrolled_college pct_
persist_college type
order first_college_name enrl_1oct_grad_yr1 pct_enrolled_college enrl_grad_persist pct_
persist_college type

gen rank = (regexm(first_college_name, "ALL"))
replace rank = 0.5 if regexm(first_college_name, "OTHER")
gsort -type rank -enrl_1oct_grad_yr1
drop rank

rename first_college_name College_Name
rename enrl_1oct_grad_yr1 Number_Enrolled
rename pct_enrolled_college Percent_Enrolled
```

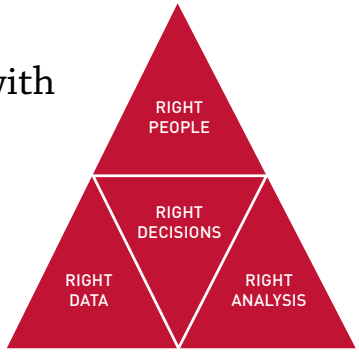

3. TOP-ENROLLING COLLEGES/UNIVERSITIES OF AGENCY GRADUATES

```
rename enrl_grad_persist Number_Persisted
rename pct_persist_college Percent_Persisted

// 6. Outsheet Table 2 (4-year colleges) into a csv file
preserve
      keep if type=="4yr"
      drop type
      outsheet using "${graphs}/E3_Top_Enrl_Col_Institutions_Table_2.csv", comma replace
restore

// 7. Outsheet Table 3 (2-year colleges) into a csv file
preserve
      keep if type=="2yr"
      drop type
      outsheet using "${graphs}/E3_Top_Enrl_Col_Institutions_Table_3.csv", comma replace
restore
}
```

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 US. **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

23 AGENCY PARTNERS
14 SCHOOL DISTRICTS
7 STATE EDUCATION DEPARTMENTS
2 CHARTER SCHOOL ORGANIZATIONS

79 FELLOWS
54 CURRENT
25 ALUMNI

CORE STRATEGIES

1. Placing and supporting top-notch analytic leaders as "Fellows" for two years with our partner agencies

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 many more education agencies

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 many more educational agencies. This toolkit is meant to help analysts in all educational 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. Later releases will address analyses relating to teacher effectiveness.

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