Colorado Department of Education CSAP

Ben Gaenslen

May 6, 2024

The Colorado Department of Education (CDE) conducted the Colorado Student Assessment Program (CSAP) consisting of standardized-tests in reading, writing, and math in years leading up to 2011. Low scores in math drew interest from the CDE and led to meaningful changes in schools' curriculum. Using data from 2009 and 2010 on the CDE website, we address the following inquiries:

- Was there truly an improvement in passing rates for the math exam between 2009 and 2010? If so, was there a decline in performance on the reading and writing exams?
- Was there an association between a school's passing rate on the math exam and its passing rates on the reading and writing exams? If so, how accurately could math passing rates have been predicted from the reading and writing rates?

Acquiring the Data

After importing data from the Colorado Department of Education API in 3 separate files, one for math, one for reading, and one for writing, we convert the file to JSON, and then into a data frame. We combine all of the the data using `rbind()`, removing rows with `school_no` equal to 0 which removes falsely identified school numbers. Then, we remove any schools with fewer than 31 total students accounted for so that outlier performances do not skew the data. Next, the data is split into two separate frames, one for each year, and counts are obtained for each type of score, `noscore`, `unsatisfactory`, `partial`, `proficient`, and `advanced`. The data is then parsed with the column `year` added.

Wrangled data

```
## Rows: 1,488
## Columns: 8
## $ school_no      <dbl> 187, 212, 221, 263, 309, 503, 15, 210, 1752, 4108, 5043~
## $ subject        <fct> MATH, MATH, MATH, MATH, MATH, MATH, MATH, MATH, MATH, M~
## $ unsatisfactory <dbl> 58, 21, 47, 32, 55, 26, 37, 28, 177, 118, 68, 222, 150,~
## $ partially      <dbl> 21, 10, 16, 5, 15, 13, 36, 10, 115, 179, 160, 150, 183,~
## $ proficient     <dbl> 14, 2, 8, 3, 4, 9, 20, 0, 48, 140, 196, 84, 150, 71, 32~
## $ advanced       <dbl> 3, 2, 0, 0, 1, 1, 4, 0, 17, 38, 88, 36, 38, 14, 4, 1, 3~
## $ not_scored     <dbl> 1, 3, 5, 0, 3, 0, 0, 5, 69, 2, 3, 16, 1, 4, 18, 1, 2, 1~
## $ year           <dbl> 2009, 2009, 2009, 2009, 2009, 2009, 2009, 2009, 2009, 2~
```

Passing rates for each school within each subject are calculated with `proficient` and `advanced` counts indicating the number of passed tests. Finally, we split the data into three separate objects, one for each subject, and added a column `diff` that represents the difference between 2010 and 2009 passing rates. The following data is associated with math, reading, and writing respectively.

Data frames for Math, Reading, and Writing

```
## Rows: 120
## Columns: 4
## $ school_no <dbl> 212, 221, 263, 503, 1752, 4108, 5043, 5816, 6060, 6402, 1021~
## $ rate09    <dbl> 0.10526316, 0.10526316, 0.07500000, 0.20408163, 0.15258216, ~
## $ rate10    <dbl> 0.11111111, 0.04166667, 0.09523810, 0.23076923, 0.05191257, ~
## $ diff      <dbl> 0.005847953, -0.063596491, 0.020238095, 0.026687598, -0.1006~

## Rows: 120
## Columns: 4
## $ school_no <dbl> 212, 221, 263, 503, 1752, 4108, 5043, 5816, 6060, 6402, 1021~
## $ rate09    <dbl> 0.6052632, 0.3636364, 0.2250000, 0.5918367, 0.5586854, 0.643~
## $ rate10    <dbl> 0.5925926, 0.3191489, 0.3571429, 0.6153846, 0.1803279, 0.727~
## $ diff      <dbl> -0.012670565, -0.044487427, 0.132142857, 0.023547881, -0.378~

## Rows: 120
## Columns: 4
## $ school_no <dbl> 212, 221, 263, 503, 1752, 4108, 5043, 5816, 6060, 6402, 1021~
## $ rate09    <dbl> 0.3421053, 0.1428571, 0.0500000, 0.5102041, 0.4154930, 0.436~
## $ rate10    <dbl> 0.25925926, 0.12765957, 0.14285714, 0.48076923, 0.11475410, ~
## $ diff      <dbl> -0.082846004, -0.015197568, 0.092857143, -0.029434851, -0.30~
```

For all 120 schools we obtain a difference in passing rates (diff) from 2009 to 2010. A negative value specifies a decrease in passing rate at that school for that subject and a positive number indicates an increase. We will now analyze the 2009 and 2010 passing rates and assess changes, as well as identify correlations between subjects' passing rates.

Data Analysis

To conclude if scores improved or regressed, we will employ a bootstrap resampling distribution, utilizing our 120 school sample, which we treat as the entire population. Then we randomly sample from our "population" with replacement and calculate 1000 means. By sampling with replacement we are able to mimic obtaining data from a larger population.

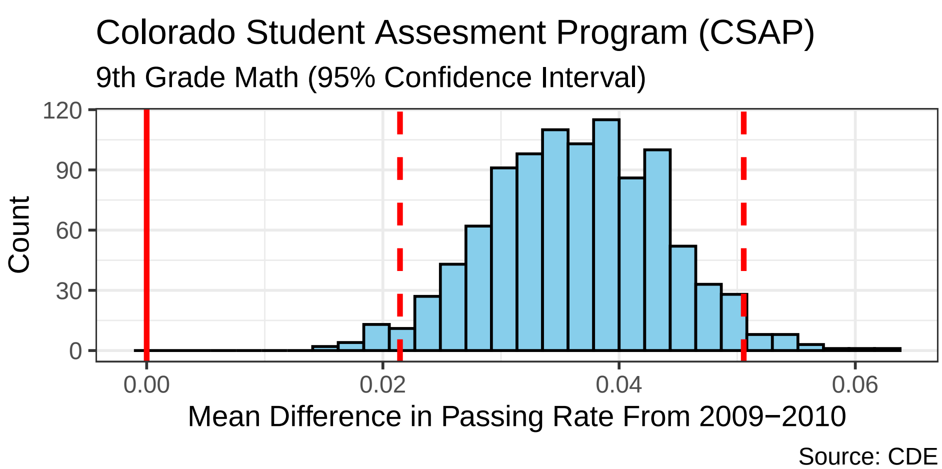

Figure 1: Mean difference in passing rate for math from 2009 to 2010

In Figure 1, the bootstrap resampling distribution with 95% confidence bounds, we can see that zero is not in the interval and thus it appears that 9th grade passing rates for math *did* increase from 2009 to 2010 for the average Colorado school. In fact, we are 95% confident that passing rates for math increased between 2.1% and 5.1% from 2009 to 2010.

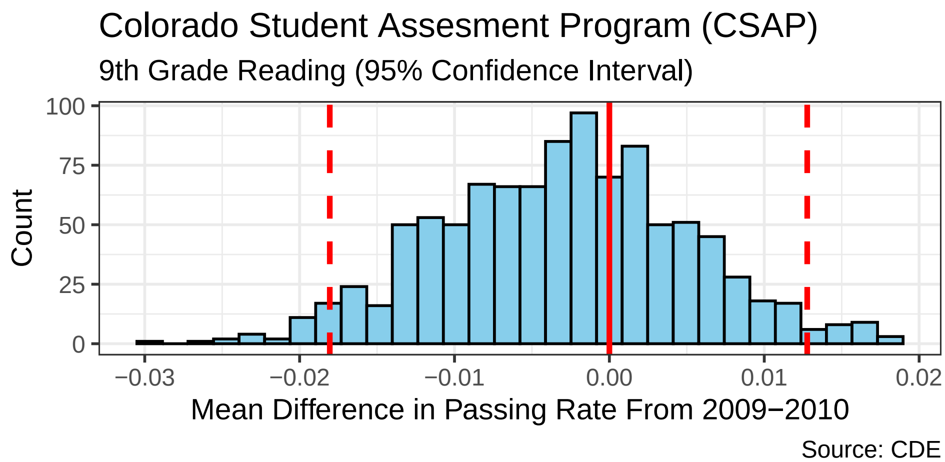

Figure 2: Mean difference in passing rate for reading from 2009 to 2010

The bootstrap resampling distribution with 95% confidence bounds in Figure 2, shows that zero *is within* the interval and thus we do not have sufficient evidence to say that 9th grade reading scores increased or decreased from 2009 to 2010 for the average Colorado school. So, we are 95% confident that the difference in passing rates for reading is between -1.8% and 1.3% from 2009 to 2010 for the average Colorado school.

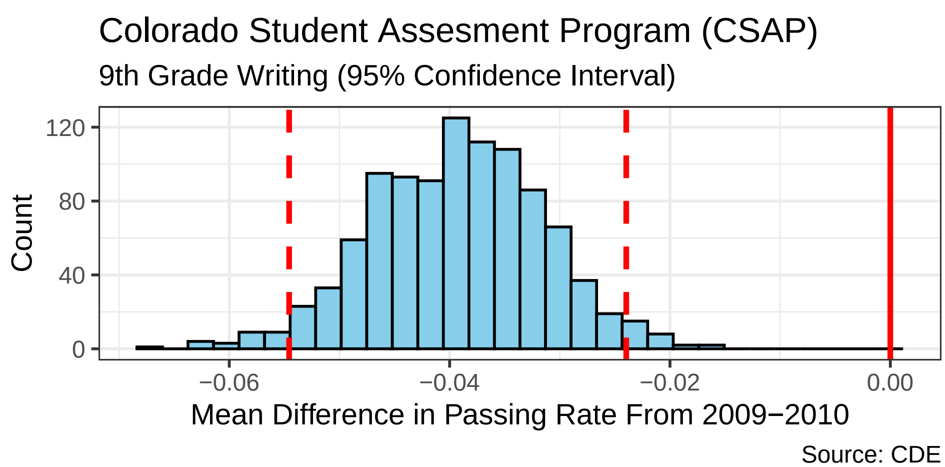

Figure 3: Mean difference in passing rate for writing from 2009 to 2010

The 95% confidence interval in Figure 3 does not contain zero and thus it appears that 9th grade passing rates for writing decreased from 2009 to 2010 for the average Colorado school. In fact, we are 95% confident that passing rates for reading decreased between 2.4% and 5.5% from 2009 to 2010 for the average Colorado school.

Now we explore the relationship between math scores and reading and writing scores. We first split the data into training and testing sets. 2009 data will be used for training and 2010 data for testing. To do this we will fit a linear model on the `training` passing rate for reading as the predictor variable for math pass rates and another model with writing as the predictor variable for math pass rates. Both models will be tested on the `testing` data set. We can visualize the relationship using scatter plots.

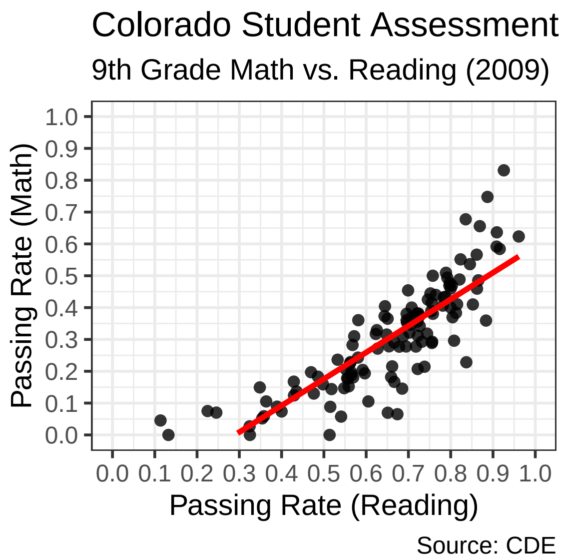

Figure 4: Association between passing rates for math and reading.

Based on Figure 4, the association between passing rates for reading and math appears linear, increasing, and moderately strong. We might expect this as schools that have higher scores in a subject would seem likely to also score higher in others as well. The correlation coefficient of 0.7 reinforces the moderate to strong correlation between reading passing rates and math passing rates. While, this is not extremely high, math and reading being different disciplines may account for some lacking predictive value of reading passing rates. The linear model has a slope of 0.835 and an intercept of -0.242, so, as the passing rate for reading increases by 1 we can expect the math passing rate to increase by 0.835 on average for schools in Colorado. The negative intercept suggests that students tend to perform better at reading than math and the slope being less than one indicates the gap in performance decreases as reading scores increase.

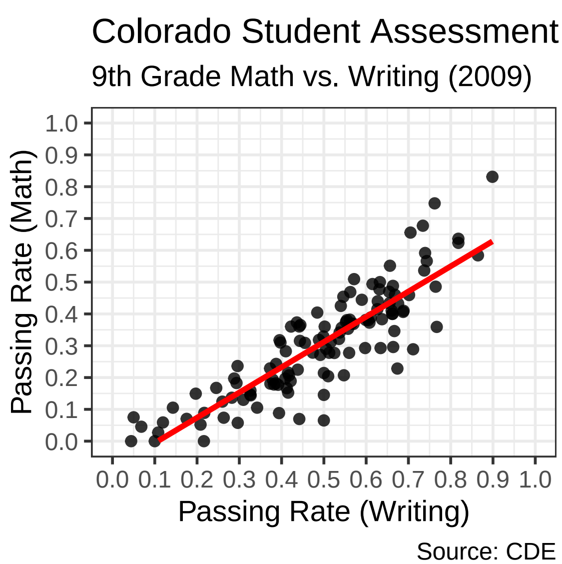

Figure 5: Association between passing rates for math and writing.

Figure 5 illustrates that the association between passing rates for writing and math appears linear, increasing, and moderately strong. We might expect this as schools that have higher scores in a subject would seem likely to also score higher in others as well. The correlation coefficient of 0.751 confirms the moderately strong correlation between writing passing rates and math passing rates. Like reading, the different disciplines may explain the imperfections of the model. The linear model has a slope of 0.793 and an intercept of -0.085, so, as the passing rate for writing increases by 1 we can expect the math passing rate to increase by 0.793 on average for schools in Colorado. The negative intercept suggests that students typically perform slightly better in writing than math and the slope being less than one indicates the gap in performance decreases as writing scores increase.

Address the Questions

Based on our confidence intervals we can infer an increase in math scores, decrease in writing scores, and maintenance of reading scores from 2009 to 2010 for 9th graders in Colorado. It is important to note that there is a 5% chance for error in our findings due to the interval we used, which may have resulted in discovering an increase, decrease or maintenance of scores where there was not. While the increased focus on math curriculum appeared to help 9th graders' passing rates, it may have diverted attention from the writing curriculum, which is perhaps why we observed a decline in those scores. It may be important to discuss what subjects are valued more than others, if at all, as well as to structure curriculum in Colorado high schools that help maintain strong passing rates in all subjects. Domain experts like the CDE should be tasked with advising these decisions, and make change accordingly.

Reading and Writing passing rates were both moderately strong in predicting the passing rates for math among Colorado Freshman, and although writing appeared to be a more accurate predictor of math passing rates given the correlation coefficient, it had a greater root mean squared error (RMSE) than the reading model. So we are able to say that given a schools reading pass rate we can predict their math pass rate to within 3.88% on average and given the writing pass rate we can predict a school's math pass rate to 6.76% on average for freshman scores in Colorado schools. Despite all the insights we obtain from the model, it has a few shortcomings. For one, or linear model did not have an extremely strong correlation coefficient, and our RMSE values may have been somewhat high. Perhaps, the difference in curriculum across the state accounted for the variance and lack of strength in our models as the model was overfit to the 2009 data.