Computer Science course enrollment and student achievement across subject areas by demographic groups

Terrie McClure
MS Data Analytics Engineering, Student
George Mason University / College of Engineering and Computing
Fairfax, Virginia, US
tmcclur@gmu.edu

***Abstract*— This study examines the relationship between high school Computer Science (CS) course enrollment and student achievement across multiple subject areas using publicly available data from the U.S. Department of Education Civil Rights Data Collection (CRDC) and the Virginia Department of Education (VDOE). School-level CS enrollment rates were merged with Virginia Standards of Learning (SOL) pass rates to explore whether exposure to CS courses is associated with higher performance in both STEM (Science, Technology, Engineering, and Math) and non-STEM subjects. Correlation and regression analyses were performed by gender and race to identify demographic patterns. Results show that higher CS enrollment is positively associated with advanced achievement across subject areas, with the strongest correlations observed in STEM tests and among Hispanic students. These findings suggest that computer science education may contribute broadly to academic success and highlight the importance of equitable access to CS opportunities for all demographic groups.**

Keywords*— *computer science, enrollment, student achievement, STEM, non-STEM, SOL pass rates, gender, race, equity

I. Introduction

Computer Science (CS) has become an essential component of modern education, yet participation rates remain uneven across gender and racial groups. Many schools have expanded access to CS courses, but enrollment data show that males continue to participate at nearly twice the rate of females (Fig 1), and representation among minority groups remains limited. Beyond career preparation, CS education can increase problem-solving, logical reasoning, and computational thinking skills that may transfer to improved performance in other academic areas.

The purpose of this research is to explore whether taking CS courses in high school is associated with higher achievement in Virginia Standards of Learning (SOL) End of Course (EOC) assessments across both STEM (Science Technology, Engineering, and Math) and non-STEM subjects.

1. Does access and participation in computer science classes in high school improve student outcomes both inside and outside of STEM?

2. Do schools with higher CS enrollment rates also have higher overall State Standards of Learning (SOL) pass rates?

3. Do schools with higher CS enrollment rates also have higher overall non-STEM SOL pass rates?

4. Does the effect of CS enrollment rates on student outcomes differ based on gender, race, or socioeconomic status?

Using combined data from the U.S. Department of Education's Civil Rights Data Collection (CRDC) and the Virginia Department of Education (VDOE), this study analyzes the relationship between CS enrollment rates and SOL pass rates by gender and race. Understanding how CS participation relates to broader student outcomes can inform policies and practices aimed at promoting equitable access and encouraging participation among underrepresented groups.

II. Literary Review

A recent study on Computer Science class enrollments in high schools shows that over the past six years, while schools are providing more opportunities to study Computer Science, enrollment by males has been steadily increasing from 13% up to 18%, but enrollment by females has remained flat at ~7% [1]. Taking Computer Science in high school could lead to better outcomes for all students. A 2024 study, "Computer Science for All? The Impact of High School Computer Science Courses on College Majors and Earnings" [3], shows an increase of approximately 8% of early career earnings by students who took Computer Science.

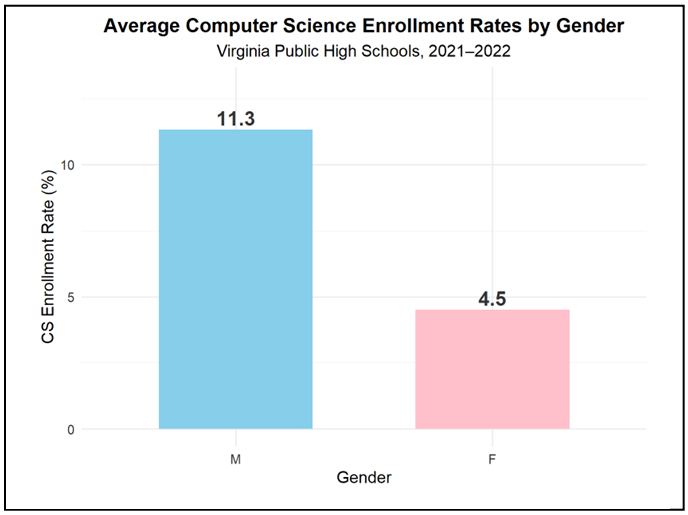

Fig. 1. Average Computer Science (CS) enrollment rates by gender in Virginia public high schools (2021–2022). Male students enroll in CS courses at more than twice the rate of female students.

The benefits of taking Computer Science can transfer to both STEM and non-STEM skills and outcomes. A meta-analysis of 37 studies by Zuokun Li and Pey Tee Oon [2] found that competency in computational thinking has positive effects on both cognitive outcomes, such as problem-solving and critical thinking, and noncognitive outcomes, including motivation and self-efficacy.

My research on the effects of computer science enrollment on SOL test pass rates in all subjects will look at whether computer science skills will transfer broadly to student outcomes in both STEM and non-STEM areas. The research looks at outcomes across race and gender. Showing the benefits of computer science education can be used to motivate students of all races and genders to enroll.

III. Methods and Limitations

Publicly available data from the U.S. Department of Education's Civil Rights Data Collection (CRDC) [8] and the Virginia Department of Education (VDOE) Statistics & Report data [10] were used to collect Computer Science course enrollment rates (CRDC) and Virginia End of Course (EOC) Standards of Learning (SOL) pass rates for public schools in Virginia. Data for middle and elementary schools were excluded. Computer science offerings are limited in middle and elementary schools, and those students who take typically high school level End of Course SOL tests (for ex. algebra) in the lower grades are usually high performing students. That data would skew the results. VDOE has a build-a-table tool that allows pulling aggregated SOL pass rates by various demographics. They have data suppression rules to protect student identities, so the more granular the data, the more likely the pass rates would be suppressed. (For example the black, female, disadvantaged, English learner, disabled population taking English reading tests at a small school would likely be suppressed.) In order to include the most schools, the SOL assessment data was pulled 3 times – 1) by summary aggregated data, 2) aggregated by gender only, 3) aggregated by race only.

The CRDC data included CS course enrollment counts by school, gender, and race. Separately, CRDC has total school enrollment counts. The two CRDC files were combined to calculate computer science enrollment rates per school. The CRDC and VDOE systems do not use the same school identifiers, and school names were inconsistent across systems (for ex. Woodson High vs. W.T. Woodson H.S.). Python code was used to clean division and school names before joining.

The combined cleaned file contained STEM and non-STEM pass rates and pass advanced rates for EOC SOL assessments as well as computer science enrollment rates by school, gender, and race. The cleaned data file was loaded into an AWS S3 data

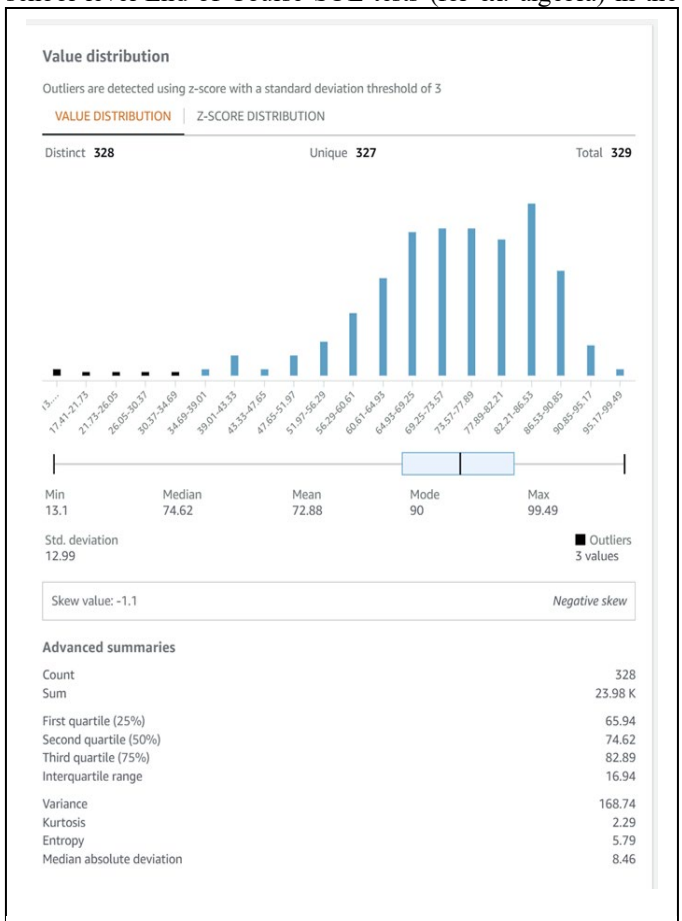

Fig. 2. Distriibution of Virginia Standards of Learning (SOL) pass rates among Virginia public high shools

Fig. 3. Distriibution of Virginia Standards of Learning (SOL) pass advanced rates among Virginia public high shools

storage bucket and was profiled using AWS Glue DataBrew data analysis tool. Characteristics of the data are shown below.

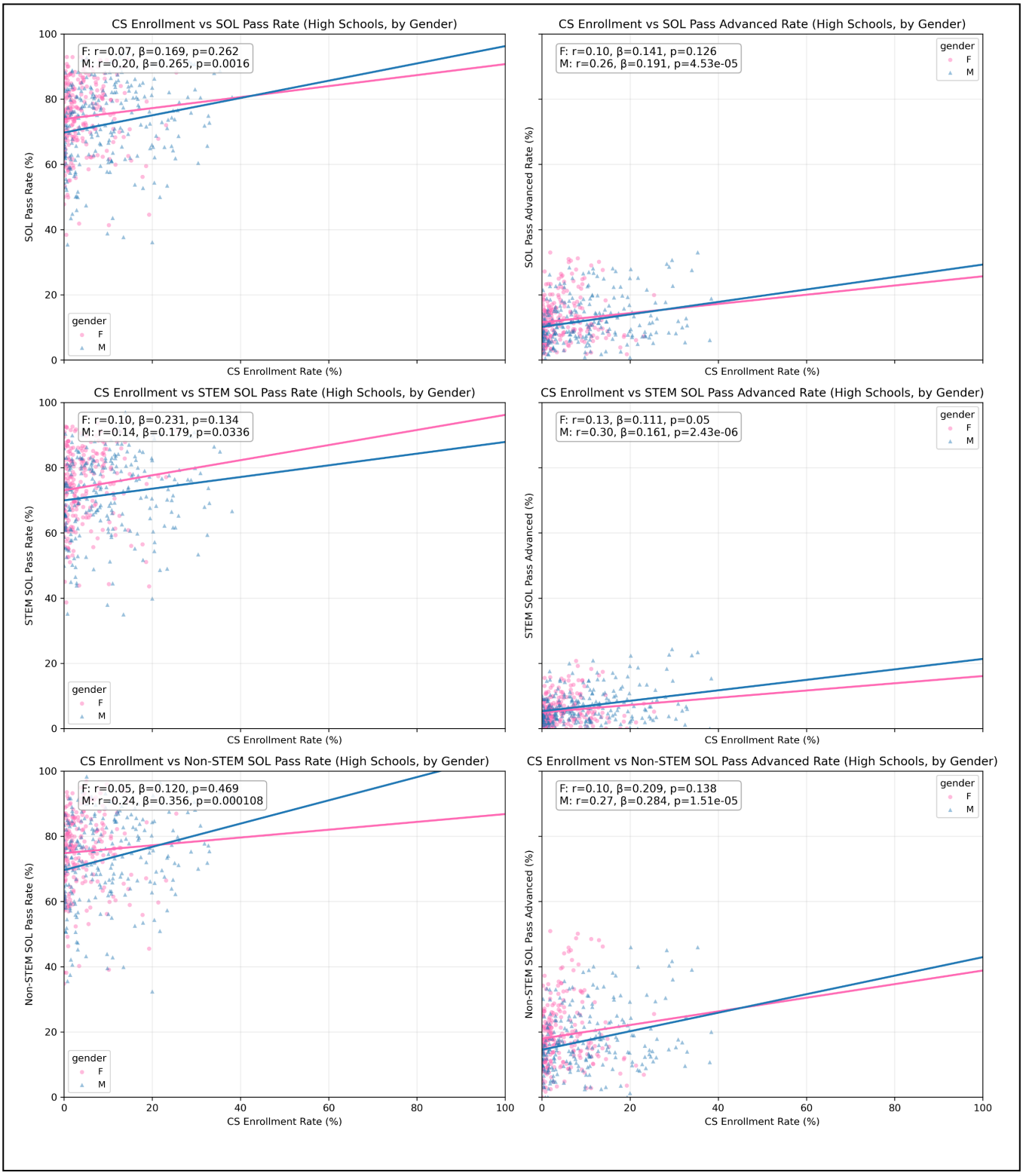

Fig. 4. Relationship between Computer Science (CS) course enrollment rate and Virginia Standards of Learning (SOL) pass rates across subject areas, by gender. Each panel displays separate regression lines and statistics (r, β, p) for female (pink) and male (blue) students. Positive correlations are similar for both groups for both STEM and Non-STEM subjects..

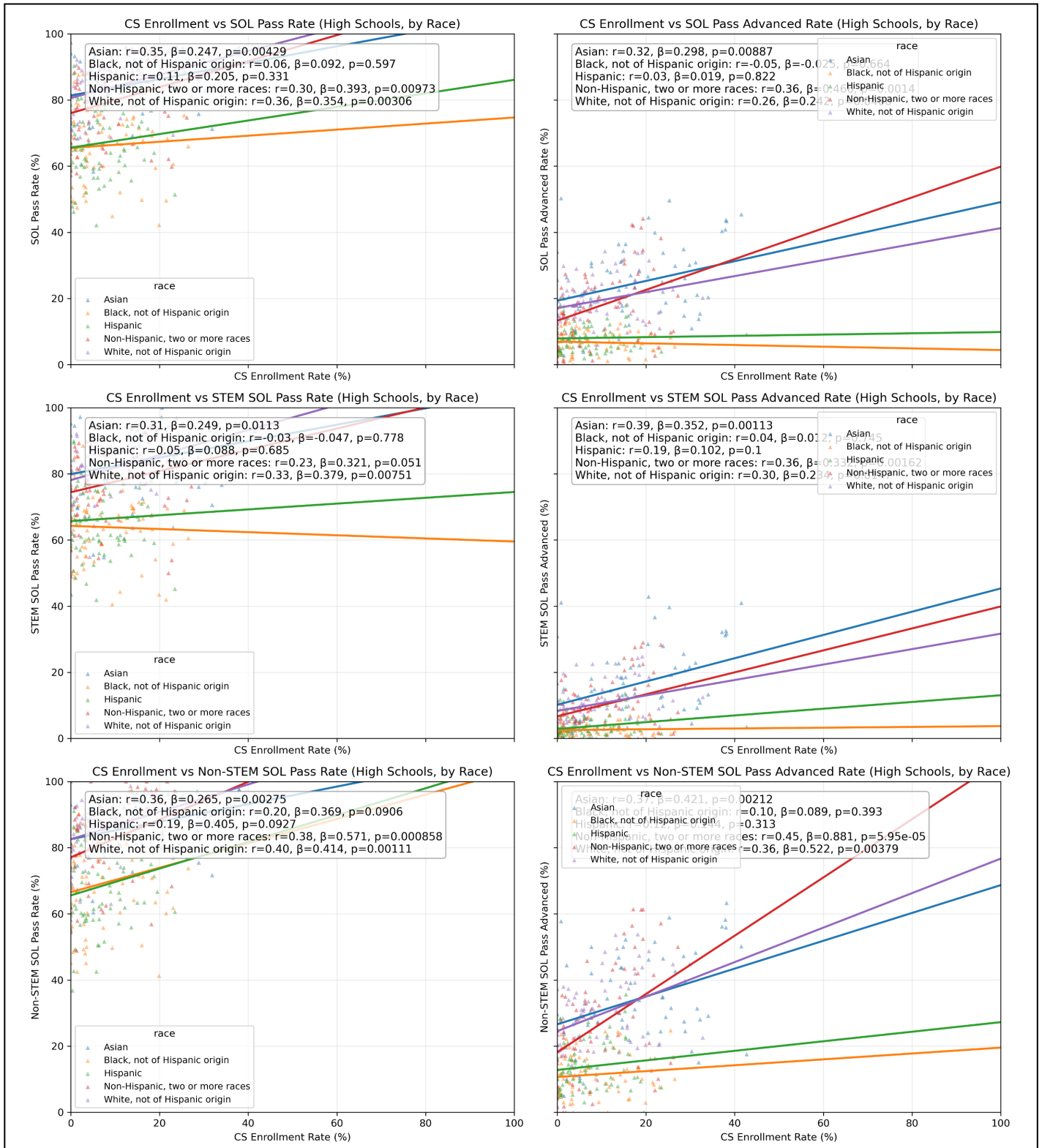

Fig. 6. Relationship between Computer Science (CS) course enrollment rate and Virginia Standards of Learning (SOL) pass rates across subject areas by race/ethnicity. Each panel displays fitted regression lines and correlation statistics (r, β, p) for Asian, Black, Hispanic, multiracial, and White student groups.

After profiling the data, outliers were removed using a Z-score method ($|z| > 3$), consistent with AWS Glue DataBrew's default profile job settings. This reduced the influence of extreme percentage values and clarified overall regression patterns across demographic groups.

Scatterplot data visualizations showing the correlation between computer science enrollment rates and SOL pass rates and pass advanced rates were generated using Python. OpenAI's ChatGPT tool [10] was used occasionally to refine the wording of explanations and to improve clarity in describing relationships shown in the data visualizations. The tool also provided limited assistance with Python code for annotating graphs with statistical information.

IV. Results and Discussion

After outlier removal, the correlation between computer-science enrollment and SOL achievement remained positive for most groups but was weak or slightly negative for Black students. This pattern likely reflects differences in school-level conditions rather than true academic effects.

Future work could incorporate additional covariates such as socioeconomic context, prior achievement, and school resources, and apply multivariate or weighted regression models to better isolate the underlying relationships.

Higher CS enrollment is positively associated with advanced performance across both STEM and non-STEM subjects. The correlation with STEM Pass Advanced rates is somewhat stronger ($r = 0.30$) (middle right, Fig 1) than with non-STEM rates ($r = 0.25$) (bottom right, Fig 1), indicating a more consistent relationship for STEM outcomes. However, the slope is steeper for non-STEM subjects ($\beta = 0.36$ vs. $\beta = 0.19$), suggesting that schools with higher CS enrollment tend to see an even greater magnitude of improvement in non-STEM advanced performance, though with greater variability. This pattern may reflect broader cognitive or problem-solving benefits of computer science education that extend beyond traditional STEM areas.

Adding race and gender to the analysis shows similar results for most demographics. Correlation and improvements to advanced pass rates are positive for both genders. For STEM advanced pass rates (middle right, Fig 2) both genders show a positive correlation. Overall pass rates in non-STEM SOL scores (bottom left, Fig 2) for girls are not as highly correlated and the increase is not as steep with CS enrollment compared to boys. This could be because girls already perform slightly better than boys in non-STEM areas leaving less room for improvement. Interestingly, overall, non-STEM improvements are greater than STEM improvements with a significant positive correlation (bottom panels, Fig 3) have the strongest correlation (r value) between CS enrollment and SOL pass rates, and show a steep slope. For example, in STEM pass advanced (top right, Fig 3), Hispanic students have $r \approx 0.70$ and a steep slope ($\beta \approx 0.70$), both quite high compared to other groups.

V. Conclusions

1. Does access and participation in computer science classes in high school improve student outcomes both inside and outside of STEM? The analysis shows that schools with higher computer science enrollment have higher EOC SOL pass rates on both STEM and non-STEM subject tests, suggesting that student outcomes are improved both inside and outside of STEM areas.

2. Do schools with higher CS enrollment rates also have higher overall State Standards of Learning (SOL) pass rates? Schools have higher overall pass rates and significantly higher pass advanced rates as CS enrollment increases.

3. Do schools with higher CS enrollment rates also have higher overall non-STEM SOL pass rates? The findings show even a steeper increase in non-STEM subject pass rates.

4. Does the effect of CS enrollment rates on student outcomes differ based on gender or race? Analysis separated by race and gender showed similar results for all demographics.

To answer Liu's question, "Computer Science for All?" [3], the study's results appear to agree that the answer is 'Yes'. Studying computer science in high school can benefit everyone – girls, boys, and any racial or cultural demographic. It can even benefit students that might be interested in subjects and future careers outside of STEM. Teachers should encourage participation in their computer science classes by everyone, especially girls and underserved populations. Girls should be motivated to enroll in computer science because it can be beneficial no matter what their future goals – it is not just for white boys aspiring to be future 'computer nerds'.

VI. Acknowledgment

This research paper was the final project for my first introductory course at George Mason University for the master's program in Data Analytics Engineering. I would like to thank my professor, Harry J. Foxwell, PhD, for creating an amazing and comprehensive course. He shared a wealth of experience and information. In 8 short weeks we covered elements of Big Data, Python and R programming, SQL, CLI, AWS Cloud including EC2, S3, RDS, Glue DataBrew, and instruction on writing research articles.

I also would like to thank my daughter, Blair McAvoy, for helping me choose a research topic by sharing her enthusiasm for excellence in high school education and curriculum for all students, high school computer science, especially for girls, and for her interest in solving education questions with data.

And lastly, I would like to thank my daughter, Shannon McAvoy, who recently completed her own master's degree in a STEM field, and offered encouragement and helpful advice on grad school and research during our daily 30 minute walks.

References

[1] Blitz, Michael, Tel Amiel, and Sam Duncan. 2025. High School Computer Science Participation: A 6-Year Enrollment Study. arXiv preprint arXiv:2503.04770. https://doi.org/10.48550/arXiv.2503.04770Li, Z., Oon, P.T. 2024. "The Transfer Effect of Computational Thinking–STEM: A Systematic Review and Meta-Analysis." International Journal of STEM Education 11 (15). https://doi.org/10.1186/s40594-024-00498-z.

[2] Liu, Jun, David Conrad, and Daniel Blazar. 2024. "Computer Science for All? The Impact of High School Computer Science Courses on College

Majors and Earnings." EdWorkingPaper No. 24-904. Annenberg Institute at Brown University. https://doi.org/10.26300/k0w5-pg15

[3] U.S. Department of Education, Civil Rights Data Collection, 2021–22 CRDC School Form, Washington, DC, 2021. [Online]. Available: https://civilrightsdata.ed.gov/assets/downloads/2021-22-crdc-school-form.pdf

[4] U.S. Department of Education, Civil Rights Data Collection, 2021–22 CRDC LEA Form, Washington, DC, 2021. [Online]. Available: https://civilrightsdata.ed.gov/assets/downloads/2021-22-crdc-lea-form.pdf

[5] U.S. Department of Education, Civil Rights Data Collection, 2021–22 Master List of Definitions, Washington, DC, 2021. [Online]. Available: https://civilrightsdata.ed.gov/assets/downloads/2021-2220List20Definitions.pdf

[6] U.S. Department of Education, Civil Rights Data Collection, 2021–22 User's Manual, Washington, DC, 2022. [Online]. Available: https://civilrightsdata.ed.gov/assets/downloads/2021-22%20User's%20Manual.pdf

[7] U.S. Department of Education, Civil Rights Data Collection, "CRDC Data," Washington, DC, 2021–22. [Online]. Available: https://civilrightsdata.ed.gov/data

[8] Virginia Department of Education, "P1PE Portal (Virginia)," [Online]. Available: https://p1pe.doe.virginia.gov/apex_captcha/home.do?apexTypeId=306. Accessed Oct. 2025.

[9] Virginia Department of Education, "SOL Test Results," in SOL Test Pass Rates & Other Results. [Online]. Available: https://www.doe.virginia.gov/data-policy-funding/data-reports/statistics-reports/sol-test-pass-rates-other-results. Accessed Oct. 2025. The Virginia Department of Education provides, via this link, an online data tool ("Build-A-Table") for the creation of customized reports on student performance on state assessments .

[10] ChatGPT (GPT-5). Accessed October, 2025. https://chat.openai.com/.