

## **REGRESSION ANALYSIS – GRADUATION RATE IN HAMILTON COUNTY SCHOOLS**

**Wolday Abrha, Ph.D., CPEM\***

**Raga Ahmed, Ph.D.**

**The University of Tennessee at Chattanooga**

**Emily McDonald**  
**Red Bank High School**

**Kevin Lee**  
**East Ridge High School**

**\*Wolday-Abrha@utc.edu**

---

### **Abstract**

Early intervention strategies for K-12 students are associated with positive educational outcomes. However, there are unique challenges facing Hamilton County schools when it comes to economically disadvantaged students (EDS) and graduation rates. This research seeks to quantify the effect of interventions and their implementation specifically in relation to EDS. It also examines if the number of EDS enrolled at a particular school had a direct impact on overall graduation rates for the 2021 graduating cohort. First, previous works on EDS and dropout rates were examined. There is plenty of relevant literature, but it became clear that there is contention among authors about the underlying reasons behind high dropout rates. Next, regression analysis was applied to analyze 2021 graduation data collected by Hamilton County on their official website for individual schools. Regression analysis of this data confirmed an earlier hypothesis but with interesting caveats. While higher numbers of EDS tend towards higher dropout rates at some schools and lower numbers tend to lower overall dropout rates at other schools, there is a lack of such correlation in schools where EDS enrollment numbers are between the two extremes. This led to a reevaluation of assumptions in favor of investigating if there is a "threshold" of students that exists at the upper and lower ends of the spectrum for enrollment in EDS or if the implementation of academic interventions differs significantly among schools. This requires further research, and a follow-up study is ongoing.

### **Keywords**

Dropout Rates, Inferential Statistics, K-12 Education, Regression, Socioeconomics

### **Introduction**

Education is an important indicator of lifetime earnings and has important implications for retirement outcomes. Additionally, a high school diploma is a standard requirement for most jobs and for higher education opportunities. The absence of a high school diploma can therefore lead to negative outcomes such as lower wages and poverty (Autor, 2014; Abdullah et al., 2015). On a bigger and global scale, deep educational crises such as equity, quality, and relevance have off-tracked the United Nations' 2030 agenda, particularly access to quality education (United Nations, 2022). On a local scale, in the United States, while education officials and various stakeholders in educational institutions are making efforts for improving access and quality of education at various levels, dropout rates continue to be unaddressed in a holistic manner. The estimated high school graduation rates in the United States have shown staggered results from state to state. The state of Montana has the highest high school graduation rate of 94% in 2023, followed by Wyoming and Vermont at 93.6% and 93.5%, respectively (World Population Review, 2023). The three states with low high school graduation rates are California at 83.9%, followed by Texas at 84.4%, and Mississippi at 85.3%.

The centerpiece of this study is Hamilton County Schools (HCS) in the state of Tennessee (TN) which has 94 school districts. The state of Tennessee is ranked 38<sup>th</sup> with a graduation rate of 88.2%. An earlier report from the National Center for Educational Statistics (NCES) indicates a slightly higher rate ranking Tennessee among the top performers at 90% or higher during the academic year 2019-2020 (NCES, 2023). The fluctuations in graduation rates in the past 4 – 5 years necessitate further analysis to investigate the underlying reasons hindering 100 percent graduation rates as well as to identify long-lasting solutions. Hamilton County is a diverse community; therefore, multiple variables contribute to graduation rates at local schools.

This study seeks to examine how the number of economically disadvantaged students (EDS) at a particular school impacts the graduation rate. The research also seeks to quantify the effectiveness of EDS interventions in improving graduation rates. A review of previous work on EDS and dropout rates reveals that there is plenty of literature on the subject, but it also uncovers contention among authors about factors (or underlying reasons) behind increased dropout rates. The major factors or reasons cited for increased dropout rates tend to fall under five main categories: disengagement, absenteeism, peers, home expectations, and grades (Azzam, 2007). An important observation here that multiple indicators under these categories could apply to a single student and that a student is more likely to drop out of high school the more indicators apply to her/him. This observation is (and these factors are) important to consider, and will be reemphasized in the Results Section, when discussing the implementation of EDS interventions.

Efforts to reduce high school dropout rates have been ongoing for quite a while, especially since the 1970s (Suh & Suh, 2007). In Suh & Suh (2007) both academic and non-academic dropout risk factors, such as low GPA, socioeconomic status, and behavioral problems are identified. However, some authors tend to (unfairly) "blame dropouts for their circumstances" by casting them as "a homogenous (often deviant) group" (Townsend et al., 2020). This paper, which belongs to a large body of research that supports continued research in the areas of high school dropout rates or graduation rates, aims to build upon the work done by Abrha et al. (2022), where the authors have developed a Lean Six Sigma-based framework to mitigate the dropout rates. That work lacked the methodology to "identify underlying inferential statistics by looking deeper into socioeconomic factors, geographical variables, and others" (Abrha et al., 2022). Therefore, this study improves and expands on that work by addressing the following research questions:

1. What is the prevalence and impact of socioeconomic factors on graduation rates?
2. How effective are interventions to improve graduation rates employed by different schools in Hamilton County? Implicitly, it is hypothesized that access to greater numbers of counselling personnel and numbers of interventions specific to helping EDS directly impact high school graduation rates.

Following the introduction section, where the description of the problem and literature review are provided, the remaining sections of the paper are arranged as follows. First, the methods are presented. Second, the data obtained from various sources are analyzed and discussed in the results and discussion section. Finally, the conclusions section summarizes the key findings of this study and outlines future research directions.

## **Methods**

First a thorough literature review and a careful examination of historical data on dropout rates in Hamilton County Schools and surrounding school districts were completed. Regression analysis was then performed on 2021 data published by Public Education Foundation and Hamilton County to produce a predictive model to inform further research toward conclusions and recommendations to the school system. Graduation rates for Hamilton County Schools between 2009 and 2017 were also used to visualize the various characteristics, such as students with disabilities, limited English proficiency, and EDS. Therefore, the primary research method for this study is qualitative based on secondary data.

Combinations of R and Microsoft Excel are used as primary tools for data filtering, analysis, and visualization, as well as for developing predictive models. Linear, exponential, logarithmic, polynomial – quadratic, polynomial – cubic, and power functions were among the trendline options explored to choose the best possible predictive model to analyze graduation rates. The value of the coefficient of determination is the main determinant employed in picking the most reasonable model. But this is done with due caution because the high value of coefficient of determination does not imply a good trend to depend on. Simplicity in computation and interpretation of the model are also equally important. Predictions using the developed logarithmic regression model are discussed in the following section.

## **Results**

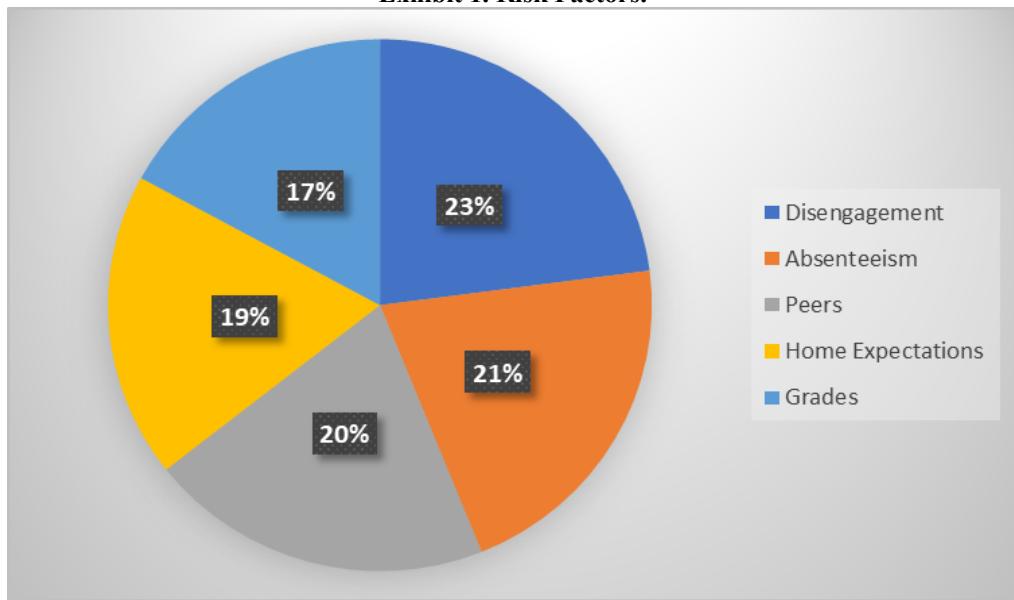
Before detailing with the regression model and analyzing other relevant information pertaining to Hamilton County Schools, it is paramount to revisit and reemphasize some of the relevant observations obtained from previous works, more specifically on why students drop out of school. The five risk factors or reasons for dropping out of school are

disengagement, absenteeism, peers, home expectations, and grades. A brief description of each of these factors is given below.

Disengagement refers to a lack of interest in school or being bored. For this case, the intervention may be to find a mechanism that motivates or incentivizes the students to stay engaged. But the disengagement can be beyond the control of the school administration. Absenteeism refers to the case when a student missed too many days and is unable to catch up with teaching/learning expectations. For example, it would not be surprising the first factor (disengagement) is found to cause the second (absenteeism). The next risk factor, peers, is when students spend time with those who are not interested in school. Home expectations refer to when students have much freedom and little or no rules in their lives. The last risk factor, grades, is purely an academic issue. It is when students fail classes.

Exhibit 1, which is based on Suh & Suh (2007), shows a typical percent contribution on dropout rates of each of the five major dropout risk factors or reasons why students drop out of school. The percentages shown are not vastly different, but the risk factors and their associated weightage or percentage can vary among schools, and it is particularly important to understand these factors to devise appropriate interventions. What is more important for the implementation of interventions or to see the response to interventions is whether these risk factors are the same or significantly different between EDS and other students. Therefore, Exhibit 1 was used as a primary input during discussion among the authors and other teachers attending a summer 2022 Research Experience for Teachers (RET) program at the University of Tennessee at Chattanooga, to learn their perception at their respective schools and Hamilton County in general.

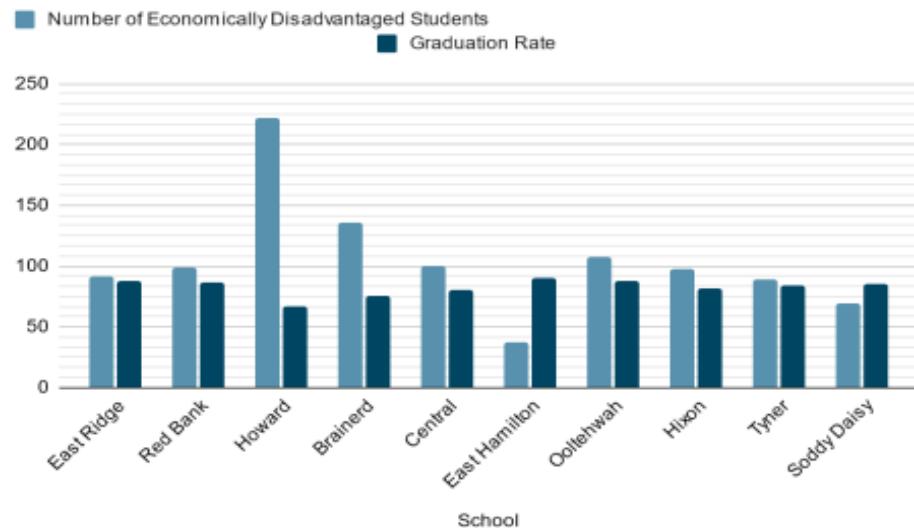
**Exhibit 1. Risk Factors.**



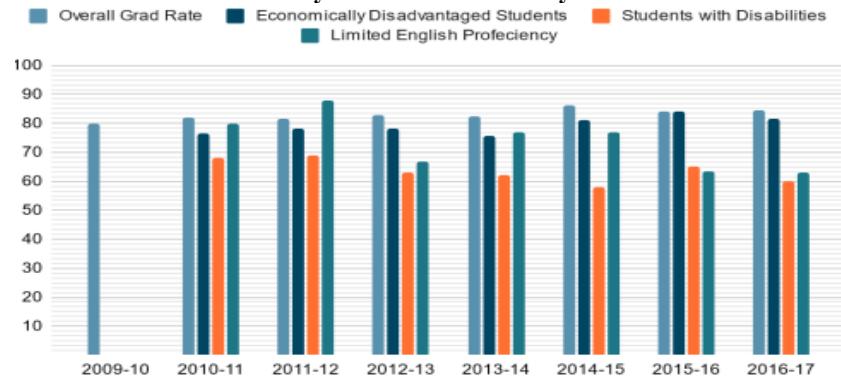
Following the discussion and general overview of the underlying reasons that may negatively impact graduation rates, the results specific to Hamilton County Schools were analyzed. Exhibit 2 compares the overall graduation rates and the number of EDS. The graphs show that three out of ten schools (Howard, Brainerd, and East Hamilton) have significant differences in the number of EDS and graduation rates, as compared to the remainder of the ten schools. In Exhibit 2, the same vertical axis is used to depict a graduation rate (as a percentage) and an EDS (a number), and so care is to be taken in interpretation and visualization. As it can be clearly seen, the graduation rates in all the ten schools included in the exhibit are less than 100 percent each.

Another observation from HCS data is that the number of EDS in all schools except Red Bank, Howard, Brainerd, Central, and Ooltewah is less than one hundred students.

The next graph, shown in Exhibit 3, is about special characteristics represented in the graduation rates from 2009 to 2017. It is noted again that in this graph the same vertical axis is used for both a percentage (graduation rate) and a number (economically disadvantaged students, students with disabilities, and limited language proficiency). The significance of the illustration of graduation rates versus the special characteristics is to show that EDS, although an important variable to study separately, is also interrelated to other variables that have their own effect on graduation rates. But of all these variables or characteristics, EDS has been the highest every year.

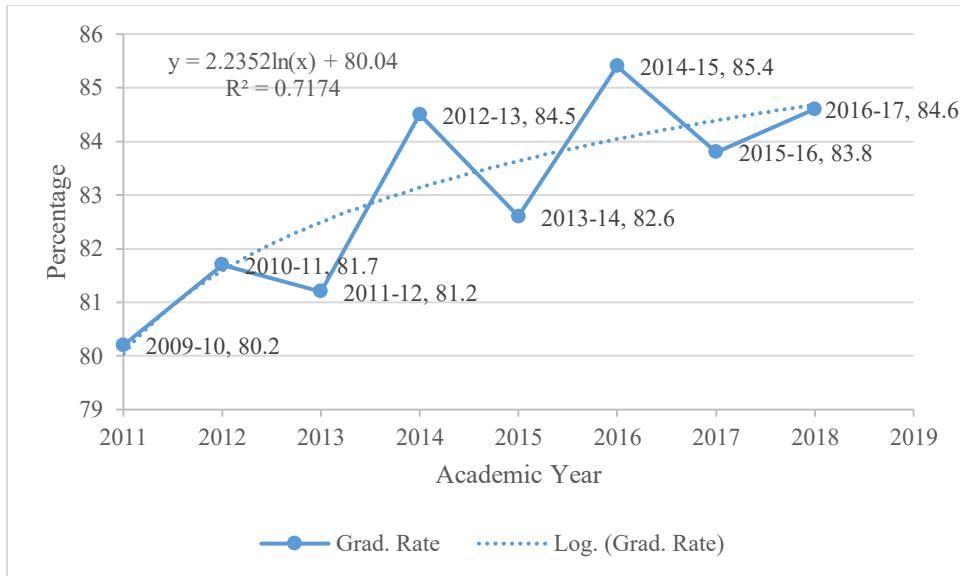
**Exhibit 2. Hamilton County Graduation Rates vs. Number of Economically Disadvantaged Students.**

The regression model predicted that the graduation rate for Hamilton County Schools for the academic year 2022-2023, is 85.77% as shown in Exhibit 4. The base year for the model is 2009, and the prediction year is therefore 13 years away from that base year. This means the independent variable  $x=13$  is plugged into the model's equation (in this case in the logarithmic equation in Exhibit 5). The same model can be used for prediction with a different base year, say taking 2020 as a base and predicting the graduation rate for 2024. Using time (year) as an independent variable should not be confused with time series analysis because the models used for prediction are purely regressive; autocorrelation, smoothing, or decomposition are not in the scope of this study. But such models and approaches may be considered in the future to provide a different perspective.

**Exhibit 3. Hamilton County Graduation Rates by Student Characteristics.**

The correlation with graduation rate of other characteristics from those described in Exhibit 3 can also be examined: Both EDS and the numbers of students with disabilities have strong correlation with graduation rate. Surprisingly, the correlation coefficient of students with disabilities of 0.79 is greater than that of EDS, which is 0.69. The correlation coefficient of limited language proficiency is 0.39. While all of these characteristics negatively impact graduation rate, one question is posed here for further analysis in the future. That is, what if a student has two or more characteristics at the same time?

To understand how well the predictive model performed or explained the observed data, based on the proportion of total variation of outcomes explained by the model, a Coefficient of Determination,  $R^2$  is used. As shown in Exhibit 5, the logarithmic model predicts well. Understandably, the dataset, 2009 to 2017 may seem like a small sample. But practically, the models predicted well. However, the larger the sample size, the better the results would be. Therefore, using a bigger dataset in the future is planned.

**Exhibit 4. Graduation Rates Prediction.****Exhibit 5. Predictive Models.**

Trend Line Options	R <sup>2</sup>	Equation
Linear	R <sup>2</sup> = 0.6663	$y = 0.6183x + 80.22$
Exponential	R <sup>2</sup> = 0.6632	$y = 80.236e^{0.0075x}$
<b>Logarithmic</b>	<b>R<sup>2</sup> = 0.7174</b>	<b><math>y = 2.2343\ln(x) + 80.04</math></b>
Polynomial (Quadratic)	R <sup>2</sup> = 0.7322	$y = -0.0972x^2 + 1.4935x + 78.761$
Polynomial (Cubic)	<b>R<sup>2</sup> = 0.734</b>	$y = -0.0085x^3 + 0.0181x^2 + 1.0533x + 79.184$
Power	R <sup>2</sup> = 0.7179	$y = 80.059x^{0.0271}$

## Conclusion

This study has built on previous studies (specifically, Abrha et al., 2022) in order to examine and analyze key risk factors in high school dropout rates in Hamilton County Schools, TN. The analysis found a positive correlation between the number of EDS in a school and corresponding dropout rate that should be interpreted with caution. The regression analysis of Hamilton County's 2021 graduation data confirmed the research hypothesis correlating an increased number of EDS with lower graduation rates. However, there is an important caveat: while the two schools with the highest and lowest numbers of EDS correspond respectively to the highest and lowest dropout rates, such correlation is not seen in schools where EDS enrollment numbers are between the two extremes. This finding led to a reevaluation of assumptions in favor of investigating if there is a "threshold" of students that exists at the upper and lower ends of the spectrum for enrollment in EDS or if the implementation of academic interventions differs significantly among schools. Such a difference in the implementation of interventions along with other limitations of the study indicates a need for further research as explained in the next section.

## Future Research

Since the study includes magnet schools, where students enrolled in these schools come from different zip codes, not necessarily zoned to the school, a separate means of analysis of interventions must be used. For instance, detailed data

broken down by zip codes, socio-economic and demographics, and availability of community engagement activities for students can help to pinpoint further data analysis. Therefore, further research must be explored with additional data to share comprehensive recommendations with the school system and generalize it as best practice for other school systems to learn from. The authors are working on the next phase of this research including a survey questionnaire to high school students, and benchmarking on response to intervention with outperforming schools.

## Acknowledgement

This paper is a product of an ongoing project sponsored by the National Science Foundation's Research Experience for Teachers (RET) program (award number: 1855242). The authors are thankful to all participants in the RET cohorts of 2021 and 2022, the principals of the participating high schools, and all partners and stakeholders.

## References

Abdullah, A., Doucouliagos, H., & Manning, E. (2015). Does education reduce income inequality? A meta-regression analysis. *Journal of Economic Surveys*, 29(2), 301-316.

Abrha, W., Gould, P., McCollum, C., & Ahmed, R. (2022). Lean six sigma for mitigating high school dropouts. In *Proceedings of the International Annual Conference of the American Society for Engineering Management*. (pp. 1-7). American Society for Engineering Management (ASEM).

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the “other 99 percent”. *Science*, 344(6186), 843-851.

Azzam, A.M. (2007, April 1). Special Report – Why Students Dropout. *Association for Supervision and Curriculum Development ASCD*, 67(7). Retrieved from <https://www.ascd.org/el/articles/why-students-drop-out>

NCES (2023, May). Public High School Graduation Rates. Retrieved from <https://nces.ed.gov/programs/coe/indicator/coi/high-school-graduation-rates>

Suh, S., & Suh, J. (2007). Risk factors and levels of risk for high school dropouts. *Professional School Counseling*, 10(3), 2156759X0701000312.

Townsend, L., Flisher, A. J., & King, G. (2007). A systematic review of the relationship between high school dropout and substance use. *Clinical child and family psychology review*, 10, 295-317.

United Nations (2022). Report on the 2022 Transforming Education Summit. Retrieved from [https://www.un.org/sites/un2.un.org/files/report\\_on\\_the\\_2022\\_transforming\\_education\\_summit.pdf](https://www.un.org/sites/un2.un.org/files/report_on_the_2022_transforming_education_summit.pdf)

World Population Review (2023). High School Graduation Rates by State 2023. Retrieved from <https://worldpopulationreview.com/state-rankings/high-school-graduation-rates-by-state>

## About the Authors

**Wolday Abrha** is a Visiting Assistant Professor at Engineering Management & Technology Department at the University of Tennessee, Chattanooga. He earned a Ph.D. in Industrial Engineering from the University of Tennessee, Knoxville. He holds MS in Engineering Technology and BS in Industrial Engineering from Middle Tennessee State University and Mekelle University, respectively. His research interests are in the areas of applying lean six sigma, and quality techniques to improve organizational performance; modeling, analyzing, and designing flexible manufacturing & supply chains; systems simulation, data analytics, and optimization.

**Raga Ahmed** is an Associate Professor at the Electrical Engineering Department at the University of Tennessee at Chattanooga. Her current research activity is focused on K-12 teachers and student support. Other areas of interest include electrical motor design and control applications.

**Emily McDonald** is a mathematics and dual enrollment statistics teacher at Red Bank High School, Hamilton County, Tennessee. She is a Ph.D. candidate at the University of Tennessee, Knoxville. Emily earned M.A. in Curriculum Instruction and B.S. in Education from Tennessee Technological University, Cookeville. She strives to foster a classroom culture where students understand mathematics in their everyday lives, promoting creativity and exploration in mathematics, and taking advantage of technology.

**Kevin Lee** teaches English at East Ridge High School, Hamilton County, Tennessee. He earned master's degree in education. from the University of Tennessee at Chattanooga and bachelor's degree in English from Austin Penny State University, respectively. Kevin has served in the United States Air Force for more six years. He relies on his military experience and involvement in social action, especially in empowering those that are marginalized, and to explain lesson materials via the lens of narrative and conflict.

Reproduced with permission of copyright owner. Further reproduction  
prohibited without permission.