5-2018

Effective Resource Utilization in Arkansas Public Schools

Ryan Sanders

Effective Resource Utilization in Arkansas Public Schools

A thesis submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Science in Industrial Engineering with Honors

by

Ryan Sanders

April 2018

University of Arkansas

Thesis Advisor: Dr. Shengfan Zhang
Thesis Reader: Dr. Art Chaovalitwongse

Signature Page

Effective Resource Utilization in Arkansas Public Schools

A thesis submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Science in Industrial Engineering with Honors

by

Ryan Sanders

April 2018
University of Arkansas

Dr. Shengfan Zhang
Thesis Director

Dr. Art Chaovalitwongse
Committee Member

Abstract

Teacher pay in Arkansas public schools varies widely from district to district across the state. This pay discrepancy is driven by both the funds available to a district and by how these funds are allocated. There is a standard per student budget given to districts across the state, but this budget can be supplemented by additional property taxes collected on property within a district. This leaves districts with more highly valued property at an advantage. Districts are free to allocate their budget for teacher pay as they see fit, with constraints on number of students per teacher and minimum teacher salary.

This research has two main objectives: 1) investigate what variables affect student performance in Arkansas public schools and 2) determine the cost-effectiveness associated with changing possible decision variables in terms of improving student performance. The objectives were achieved by using public data available through the Arkansas Department of Education. Objective 1 was accomplished using feature selection and predictive modeling. Objective 2 integrated the results found from the first objective with district budget information in order to analyze the cost-effectiveness of different district budget policies. Results from this study are valuable to districts trying to improve student performance in the most cost-effective way.

Acknowledgements

I would like to thank Dr. Shengfan Zhang for all the work she has put into me over the past few years. She has been a great advisor, but more importantly she has been my mentor and friend. I would also like to thank Dr. Art Chaovalitwongse for helping me along the way, along with Dr. Manuel Rossetti, who sparked my interest in industrial engineering and research freshman year.

My father, Jack Sanders, sparked my interest in this topic. Both he and Mr. Gary Anderson were very helpful and patient in explaining the history of education funding in Arkansas to me. I would also like to acknowledge Dr. Gary Ritter and Dr. Robert M. Costrell in the Department of Education Reform for helping me narrow down the exact question this study attempts to answer.

This would not have been possible without the help of an Honors College Research Grant from the University of Arkansas and a SIIRE grant sponsored by NSF.

Table of Contents

1. Introduction

There is a massive pay gap in teacher salaries in public schools across the state of Arkansas. The first-year teacher salaries range from the state minimum of $31,000 to $47,000 per year depending on the school district they teach in (Teacher Salary Schedule Analysis 2016-2017). This inevitably draws higher teacher supply to schools with high salaries and leaves schools in districts with lower salaries at a disadvantage. On top of the talent discrepancy across districts, low teacher pay is the main cause for a large number of teachers leaving the profession each year. According to a recent study on teacher retention in Arkansas, 30.6% of all teachers in Arkansas leave for another career within three years, and over half of all teachers in the state say that higher salaries/better benefits would keep them in the profession (Arkansas Bureau of Legislative Research 2016).

Individual teacher pay is mainly based on the number of years of teaching experience and level of education, as reflected by their degrees and the number of certifications they have (Teacher Salary Schedule Analysis 2016). Districts across the state are allotted a standard dollar amount per student as their base budget. This is a result of the 2002 Arkansas Supreme Court ruling that declared school districts must receive the same base funding per student (Lake View School District v. Huckabee, 2002). However, this amount can be supplemented by increasing property tax in a school district, known as a millage, provided that the increase in funds has a designated purpose. Districts in a more populated area are at an advantage since an increase in the millage corresponds to a much larger amount of revenue. These increased funds enable districts in more populated areas to have a larger percentage of their budget available to pay and retain teachers (2015-2016 Annual Statistical Report).

Districts can allocate resources according to their goals and priorities. Allocation of these funds results in a trade-off for districts that can be simplified into three options. First, districts can pay teachers more in order to attract better teachers and reduce turnover. Higher teacher salaries result in better quality teachers with more experience due to reduced turnover (Hendricks, 2013; Papay & Kraft, 2015). Secondly, districts can hire as many teachers as possible in order to achieve a lower student-teacher ratio. Lower student-teacher ratio results in higher student achievement (Krueger, 2003; Gilpin & Bekkerman, 2012; Chingos, 2013). Lastly, districts can also use funds to invest in special projects that may impact student achievement. Most of the time districts choose a combination of these three policy options.

Another option to be considered is the possibility of consolidating two smaller schools into a larger one. Districts are often consolidated as a way to increase the cost effectiveness of district resources. This technique has been a controversial topic in Arkansas for a very long time (Barnett et al., 2004; Bleed & Wickline, 2006). At a school level as well, combining two smaller schools with a district to a larger one can reduce the cost per student. Opponents of consolidation claim consolidation hurts both the students and the community. They claim students get more attention and are more engaged in smaller schools. Additionally, they claim that when smaller schools are closed, the corresponding communities surrounding the schools lose the center of public life (Nelson, 1985). Proponents of consolidation say the proportionate saving in costs gained by an increased level of students, or economies of scale, are worth the possible detriment to the community and students, in fact, receive better education in larger schools (DeYoung & Howley, 1992; Office for Education Policy, 2010).

In order to increase the quality of education for Arkansas students, it is important to understand what factors will most cost-effectively improve student achievement. There should not be a difference in the quality of public education in different parts of the state, yet there is a disparity in education due in part to resource availability as limited by district location. Different programs have been tested to try and decrease the gap in education quality across the state (Barth & Nitta, 2008). In order for any program to be as effective as possible, there is a need for research to identify what controllable factors drive student achievement.

This research seeks to explore cost-effective ways of utilizing resources to improve student performance while considering characteristics associated with school districts. Specifically, this research aims to (1) perform statistical analysis to examine the effect of district budget, school size, discrepancy in teacher pay, and student-teacher ratio on student outcomes in Arkansas public schools, adjusting for demographics and other attributes of school districts, and (2) conduct a cost-effectiveness analysis to evaluate various budget allocation policies in terms of the tradeoffs between total expenditure and improvement in student performance.

The remainder of this thesis is organized as follows. First, data collection and preparation will be discussed. This section will include data acquisition, definition of all variables used, and an explanation of how the data was prepared to be used in modeling. Next, methodology is described for feature selection and predictive modeling, and how the results from the predictive modeling were used to perform cost analysis. Finally, all results from this study will be presented and then discussed.

2. Data Source and Preparation

All data used came from two government sources that are publically available. All district and school information came from the Arkansas Department of Education (ADE) website (ADE My School Info n.d.) and all county information came from United States Census Bureau (USCB) estimations (United States Census Bureau n.d.). This study used data for three school years (Fall 2013 - Spring 2016), which is all that is currently available online. In the following sections, performance measures and contributing factors considered in this study are introduced, and necessary data processing steps are discussed.

2.1 Performance Measure

There are several performance metrics that can be used to evaluate a student's achievement, but to simplify this study only test scores were used in this study as the performance measure (or the response variable). Due to inconsistences in school format across Arkansas high schools (Arkansas Department of Education n.d.) and lack of standardized tests besides the ACT, high schools were excluded. Standardized tests are given each year for Grades 3-8 in every public school, so this research focuses on schools that contain these grades. Arkansas has adopted a new standardized test each of the last three school years (Arkansas Department of Education n.d.), but along with other subjects, each grade was tested for math and literacy each year. The dataset was narrowed down to only schools which contained at least one grade in Grades 3-8. Since math and literacy were tested each year, only scores from those sections of the exam were considered. In order to have a single response for each instance of data, a weighted test score was calculated for each school in that school year. The weights for each grade were calculated as the percent of the students enrolled in that grade out of the total students enrolled in Grades 3-8 in that school. The average test score for each school was then calculated using

$$School\ Test\ Score = \sum_{g=3}^{8} w_g(M_g + L_g)$$

where w_g = Percent of students in grade g in the school; M_g = Average math test score for students in grade g; and L_g = Average literacy test score for students in grade g.

Since a different test was given each year, test scores had different scales and distributions. This was accounted for by applying a Box-Cox transformation to the school test scores for each year. The optimal lambda value for each year's Box-Cox transformation was applied, the results of which can be found in the appendix (Figure 1). Applying the Box-Cox transformation gave each year's test scores a Gaussian distribution. Transformed test scores were then scaled to have a mean of 0 and standard deviation of 1 allowing for test scores from different years to be compared, despite different tests having been given (Figure 2).

2.2 Contributing Factors

There are two types of contributing factors that may influence student performance. They are classified as either decision variables or other explanatory variables. Decision variables are those that can be controlled from either a district or state allocation perspective. Other explanatory variables are used to explain the rest of the variance in the response.

In this study, decision variables include student-teacher ratio (Chingos, 2013), years of teacher experience (Papay & Kraft, 2015), teacher pay (Hendricks, 2013), school size (Nitta et al., 2010), and funds available to the district (Tow, 2006). Teacher pay can be broken down into starting salary, average salary, and average salary increase per year. District funding includes revenue streams and expenditures and is broken down into many different classifications.

Money invested in projects unrelated to teachers, such as building improvements and extracurricular, could play a significant role in student performance as well.

Variables that are not considered decision variables will be used to explain the rest of the variation in the response. Student performance is highly correlated with demographic and socioeconomic factors (Hanushek, 1997). It is important to accurately reflect the impact of resource utilization (i.e., teacher pay, district budget, student-teacher ratio) on student achievement adjusting for these other factors. Demographic features include any possible student classifications such as race and special statuses. Socioeconomic features come at the district and county level, and include variables such as median income, occupation, and property values. All 126 variables identified and their descriptions can be found in the appendix (Table 6).

2.3 Data Preparation

Data was cleaned to have the correct format for numbers and text. Some schools had to be excluded from the dataset. Initially, the dataset contained 2611 instances where a school had students enrolled in at least one Grade 3-8. In 220 instances the school was not an ordinary public school, leading to null values in the data. In 23 instances the school was classified as an alternative school. 69 schools were either closed or opened during the 3-year span. 3 instances did not have enough students to report test score. All of these instances were removed from the dataset. In all of these cases, if any school had to be removed for at least 1 year, all 3 years of data were removed. This was done to keep consistency in the data being tested, as well as to help with data analysis done later on. The dataset was reduced to 760 schools with all 3 years of data for a total of 2280 instances.

Some variables of interest in this study were not given directly and had to be calculated. First, average teacher salary increase per year was calculated using the salary schedule for each district (Arkansas Department of Education Salary Reports, 2017). Salaries increase linearly for the most part, so average teacher salary increase was said to be the salary at the maximum experience level listed (usually 15 years) minus the minimum experience level listed (0 years) and divided by the number of years on the schedule. Second, county level data from the USCB website were estimates based on the calendar year. To get estimates for the school year, the corresponding years were averaged. For example, for the 2013-2014 school year, county estimates from 2013 and 2014 were averaged together. These variables included mean and median income as well as occupation information. Lastly, variables regarding district revenues and expenses are only shown as totals for the district. To accurately compare districts, these variables were divided by the total number of students in the district. The same was done to school level variables that were population totals and not already percentages. By dividing these variables by the number of students, schools can then be compared to one another.

Some of the county variables were percentages of total population in several categories that each add up to 100%. This was a problem since all variables in each group would be highly negatively correlated. There were three different groups of these variables, totaling 22 variables, all regarding occupation: "Class of worker", "Industry", and "Occupation". In order to reduce dimensionality, k-means clustering was used (Zaki & Meira, 2014, p. 333). To find the best number of clusters, sum of squared errors, silhouette score, and Calinski-Harabaz score (Zaki & Meira, 2014, p. 450) were each plotted against the number of clusters. The "elbow" of each of these graphs was at 3, indicating that using more than 3 clusters had diminishing returns. Schools were assigned a county occupation cluster based on the 3 clusters found. It was found that there is a statistically significant difference in test scores of each of the 3 clusters. DBScan algorithm (Zaki & Meira, 2014, p. 375) was also investigated, but for no parameters was it superior to k-means clustering. Results from clustering can be found in the appendix (Figure 3, Table 7).

3 Modeling and Analysis

This section presents the main methodologies used in this study for our research objectives. First introduced are the procedures for selecting variables of interest. In the predictive modeling section, different regression methods are used in order to find which is the best at predicting student test scores. Lastly, the cost analysis section focuses on calculating the cost of increasing student test scores by one standard deviation using each of the identified decision variables.

3.1 Feature Selection

Analysis of this dataset was done using Pandas, Sklearn, Scipy, and Matplotlib packages in Python (https://www.python.org/). The processed data from Section 2 included 126 variables. It was found that many of the variables were highly correlated. Variables with high correlations to other variables and a high variance inflation factor (VIF) when performing a multiple linear regression were identified as candidates to be removed. In order to account for multicollinearity and reduce the dimensionality of the dataset, both penalized and stepwise regression models were investigated (Jha et al., 2017).

Both a lasso regression and stepwise regression were performed on the data in order to identify which variables were not important to the model. Lasso regression models penalizes variables that do not add as much information to the model by putting a limit on the sum of the absolute values of the coefficients (Fonti & Belitser, 2017). Variables that have their coefficient set to 0 are not valuable to the model. The lasso regression was performed with different values of L1 ranging from 0.01 to 0.2. In stepwise regression, variables are chosen through forward and backwards selection in order to enhance the model (Zhang, 2016). Similar to the lasso regression, if a variable does not significantly contribute to the model, the coefficient of that variable is set to 0. Variables that were previously identified as being highly correlated were removed from the dataset if the corresponding coefficients from both the lasso and stepwise regression models are zero (Fonti & Belitser, 2017). The dataset was reduced from 126 to 27 features using this method. These 27 features with descriptive statistics can be found in the results section (Table 1).

In order to more accurately calculate the effect each variable has on the response, interaction terms needed to be added. All data was first centered and standardized, and then multiplied to get the interaction terms between each variable. Variables were centered and standardized before creating interaction terms so that the coefficients from the regression model would be easier to interpret (Schielzeth, 2010; Enders & Tofighi, 2007). These second order interaction terms were added to the new dataset and a lasso regression was performed again. Interaction terms of a higher degree were not considered due to both the size of the dataset and the fact that higher degree interactions are usually statistically insignificant. A separate lasso regression model was built for each year of data to ensure that all three years produced similar models. As done previously, the variables with a coefficient of 0 that were not decision variables and exhibited a high degree of correlation with another variable were removed from the dataset. This added a total of 29 interaction terms to the dataset (Table 8).

3.2 Predictive modeling

By exploiting the idiosyncratic variation across districts, the effect of the decision variables can be measured. Regression analysis (Draper & Smith, 1998) was performed to identify the factors that are significantly associated with student performance improvement.

Different regression models were tested to see which would best predict student performance. The following regression methods were identified to model the dataset: multivariate linear least-squares, multivariate linear least-angle, elastic net, pure lasso, pure ridge, MARS, ARD, Bayesian ridge, orthogonal matching pursuit, and kernel ridge. These models were chosen both because of their widespread use in literature as well as their ease of implementation using the sklearn package in python (Scikit-learn). For each of the models that accept different parameters, a wide range of parameters was used to identify the best values for the parameters. A standard scaler was used to scale each of the variables separately for each year.

Each identified model was applied to one year of data for training and another year of data for validation. Models were cross validated by using each combination of the 3 years to train and test for a total of 6 runs. Metrics for model evaluation include R^2, explained variance score, and mean squared error, and run time. The average of the values for each metric from the training sets are shown in the results section (Table 2).

3.3 Cost Analysis

The cost-effectiveness of different district policies can be determined using available district budget information along with the results of the regression analysis. Using the linear regression model derived in the first objective, the cost-effectiveness of different policies was calculated.

Multiple linear regression was chosen because the coefficients given from this model are easy to interpret.

The coefficients of the best regression model were used to calculate the cost of different budgetary policies. For each variable, the mean of the coefficient from each year of cross-validation was found. The following decision variables were found to have an impact on student performance: average teacher salary, student teacher ratio, classified staff ratio, restricted revenue from the state, compensatory education expenditures, percent of teachers with master's or advance degrees, average years of teacher experience.

Since these variables have been standardized to a normal distribution, the coefficient of each variable can be interpreted as the amount of standard deviation change in test scores corresponding to one standard deviation increase in the variable. In other words, the actual cost of increasing the test scores of a school one standard deviation ($C_{improvement}$) can be determined via the following equation.

$$C_{improvement} = C_{dv}/coef_{dv},$$

where C_{dv} is the cost of increasing the decision variable by one standard deviation; $coef_{dv}$ is the coefficient of the decision variable in the linear regression model.

The cost per school to raise teacher salary one standard deviation is calculated by multiplying the standard deviation of teacher salary by the number of teachers per school. The cost per school to raise restricted revenue from the state is calculated by multiplying one standard deviation of restricted revenue per student by the number of students per school. The cost per school to raise compensatory education expenditures is calculated by multiplying one standard deviation of compensatory education expenditures per student by the number of students per school. The cost per school of increasing average years of teacher experience is calculated by multiplying one standard deviation of average years of teacher experience by average number of teachers per school then by average salary increase per year. The savings per school associated with raising the student teacher ratio is the inverse of the standard deviation of student teacher ratio multiplied by average number of teachers per school then by average teacher salary.

In a report by Chingos (2011) it was calculated that for the average school in the US, an increase in average class size by 5 students would result in an across the board increase of 34% in teacher salaries if all savings were devoted to that purpose. The same calculations were performed for the dataset used for this research.

4. Results

Descriptive statistics for the 27 variables selected are shown in Table 1. Of all the regression models tested, linear, Bayesian ridge, and kernel ridge performed the best. The average scores for each performance measures across all training sets are shown in Table 2. The residual plot (Figure 4) and coefficients from this model (Table 9) can be found in the appendix.

Table 1. Selected features and descriptive statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Normalized Test Score	0.000	1.000	-2.728	3.050
County Occupation Cluster	1.421	0.718	0.000	2.000
Isolated Status	0.099	0.298	0.000	1.000
County Population	80057.6	88666.5	4337.5	308102.5
Average Commute Time	21.948	3.586	15.000	38.050
Food Stamps	15.144	4.738	7.250	37.350
Median Income (Families)	50811.8	8177.3	32919.5	67296.5
Property Tax Revenue	3918.9	2162.7	951.7	19710.3
Compensatory Education Expenditures	315.3	184.4	16.4	1699.2
Facilities Expenditures	858.2	1493.0	0.0	13064.8
State Restricted Revenue	1300.5	850.8	347.4	6423.7
Unrestricted Revenue	9004.2	1244.4	7139.6	20786.8
Average Salary	47292.3	6162.2	32611.4	60336.1
Student Teacher Ratio (Calculated)	12.389	3.395	2.432	22.682
Attendance Rate	0.947	0.016	0.785	1.000
Other Race Percent	0.044	0.044	0.000	0.328
Black Percent	0.185	0.262	0.000	0.997
Hispanic Percent	0.106	0.140	0.000	0.822
Foster Percent	0.003	0.005	0.000	0.049
Male Percent	0.515	0.030	0.391	0.639
Special Education Percent	0.121	0.036	0.015	0.264
Total Students	422.7	200.5	69.0	1799.0
Disciplinary Actions Ratio	0.495	0.690	0.000	7.261
Free/Reduced Lunch Percent	0.668	0.192	0.070	1.000
Classified Staff Ratio	0.041	0.024	0.000	0.242
Advance Degree	0.396	0.139	0.029	0.853
Average Years Experience	12.1	3.4	1.4	23.9
Completely Certified	0.986	0.026	0.792	1.000

Table 2. Averages values of model performance metrics from training sets

Model	R^2 Mean	R^2 St Err	EVS	MSE	Time
LinearRegression()	0.6545	0.0136	0.6545	0.3455	0.0018
ElasticNet(alpha=0.05, l1_ratio=0.5)	0.6167	0.0151	0.6167	0.3833	0.0003
Earth()	0.6706	0.0176	0.6706	0.3294	6.2734
ARDRegression()	0.6473	0.0140	0.6473	0.3527	4.2916
BayesianRidge()	0.6467	0.0144	0.6467	0.3533	0.0052
Lars()	0.4821	0.1192	0.4821	0.5179	0.0104
OrthogonalMatchingPursuit()	0.5345	0.0186	0.5345	0.4655	0.0000
Ridge(alpha=1.0)	0.6544	0.0136	0.6544	0.3456	0.0000
Lasso(alpha=0.05)	0.5780	0.0167	0.5780	0.4220	0.0000
KernelRidge(alpha=1)	0.6544	0.0136	0.6544	0.3456	0.0403

The results for the cost analysis is shown in Table 3. For each decision variable the cost of increasing test scores by one standard deviation solely by increasing the decision variable is presented in the third column. It was found that the most cost-effective of improving test scores is by spending more money on teacher salary. It would cost $1,193,500 per school to improve test scores one standard deviation if all that money was spent solely on increasing teacher salary.

Table 3. Cost analysis results for decision variables

Variable	Coefficient	Std Dev	Cost per school per st dev	Benefit per $1000
Average teacher salary	0.177105872	6162.182779	$ 1,193,487.70	0.0008379
Student teacher ratio	0.161037306	3.394897491	$ (36,565,509.45)	-0.0000273
Restricted revenue from the state	0.042843371	850.7613642	$ 8,393,800.12	0.0001191
Compensatory education expenditures	0.006973674	184.4045161	$ 5,302,466.82	0.0001886
Percent of teachers with master's or advance degrees	-0.00970091	0.138928667	$ (1,938,251.27)	-0.0005159
Average years of teacher experience	0.033918523	3.353561421	$ 2,172,818.73	0.0004602

The savings associated with increasing student-teacher ratio by different amounts was calculated and the shown in terms of average teacher salary. By increasing the student-teacher ratio by 5 students, schools could increase average teacher salary by over 40%. According to the model, this would increase test scores by 0.55 standard deviations. Results from this analysis are shown in Table 4 for different values of increasing the student-teacher ratio.

Table 4. Results for increasing student teacher ratio (STR)

STR increase	New STR	Avg Teachers per School	Salary increase per teacher	Salary % increase	Std Dev Test Score Increase
0	12.323	34.30175439	$ -	0.00%	0.00
1	13.323	31.7271369	$ 3,837.71	8.11%	0.11
2	14.323	29.51202632	$ 7,675.42	16.23%	0.22
3	15.323	27.58603713	$ 11,513.13	24.34%	0.33
4	16.323	25.89603195	$ 15,350.84	32.46%	0.44
5	17.323	24.40114313	$ 19,188.55	40.57%	0.55

5. Discussion & Conclusion

The implications of this research can be used to drive policy at a district level, as well as a state level. The most cost-effective ways for districts to increase student performance are to (1) increase average teacher salary, and (2) increase average years of teacher experience. Both of these two methods had been identified in literature as being correlated and both effective ways to increase teacher quality and increase student performance (Hendricks, 2013; Papay & Kraft, 2015). Furthermore, districts should consider increasing student-teacher ratio and applying the resulting savings toward teacher salaries.

It was found that increasing student-teacher ratio was actually beneficial to test scores using this dataset. This may be due to a lack of explanatory variables in the data, as many studies have shown that lowering the student teacher ratio has a positive impact on student performance (Krueger, 2003; Gilpin & Bekkerman, 2012; Chingos, 2013). There are many variables, such as teacher effectiveness, that could not be found and which may have a large interaction effect with student-teacher ratio. In addition, only around 60% of the variation in the response can be explained by the independent variables. This is further proof that not enough variables are available to be able to accurately model the effect of student-teacher ratio on test scores. Another factor that might be causing this could be the complicated budget allocation to school districts. Some of the funding categories mentioned in the resource allocation plan (Arkansas Bureau of Legislative Research 2018) could not be found broken down in the data available on the ADE website. Yet another explanation could be that schools with more teachers higher poorer quality applicants, or that accountability for teachers at these schools are lower. This would make sense in an environment where teachers are plentiful and therefore do not have to be individually effective as teachers in an understaffed school would be. The effectiveness of teachers is not a variable accounted for in our model.

Another limitation of this study is the response variable. High schools were excluded, which greatly reduced the number of data points. There are other measures such as attendance, discipline, and high school and college graduation rates that are proxy measures of student achievement and could be used as the response variable in future research. There was also a failure to account for bias of each test toward different demographics since the standardized test was changed each year for the years available. For instance, students in an urban area may be more likely to have done well on one year's test and students in rural areas may have been more likely to do well on another year's test. This study assumes that each year's test was made fairly without bias to any population over another.

In the future, data mining methods other than regression such association and classification methods (Han et al., 2011) will be explored to further understand the relationship between resource utilization and student performance. Time varying coefficient models (Fan & Zhang, 2008; Wang et al., 2008) can also be used in order to get more insight into how policy changes have impacted students. In order to further understand why student teacher ratio actually causes test scores to increase in this dataset different stratification methods will be investigated as well as adding quadratic and higher level interaction terms.

6. References

Arkansas Bureau of Legislative Research, The House and Senate Interim Committee on Education. (2016). Teacher Recruitment & Retention Research Report.

Arkansas Bureau of Legislative Research, The House and Senate Interim Committee on Education. (2018). The Resource Allocation of Foundation Funding for Arkansas School Districts and Open-Enrollment Charter Schools.

Arkansas Department of Education, Financial and Administrative Services. (2017, March 21). Teacher Salary Schedule Analysis 2016-2017. Retrieved from http://www.arkansased.gov/divisions/fiscal-and-administrative-services/publication-and-reports/report_categories/salary-reports

Arkansas Department of Education, Financial and Administrative Services. (2017, February 13). 2015-2016 Annual Statistical Report. Retrieved from http://www.arkansased.gov/divisions/fiscal-and-administrative-services/publication-and-reports/report_categories/annual-statistical-reports

Arkansas Department of Education Salary Reports. (2017, March 21). Retrieved from http://www.arkansased.gov/divisions/fiscal-and-administrative-services/publication-and-reports/report_categories/salary-reports

ADE My School Info; Generate Statewide Report. (2017). Retrieved October, 2017, from https://myschoolinfo.arkansas.gov/

Arkansas Department of Education. (2017). Retrieved October, 2017, from http://www.arkansased.gov/

Barnett, J., Ritter, G. W., & Lucas, C. J. (2004). Does Size Matter? School Consolidation Policy Issues in Arkansas. Arkansas Education Report, 1(2).

Barth, J., & Nitta, K. (2008, February). Education in the Post-Lake View Era: What is Arkansas Doing to Close the Achievement Gap (Rep.).

Bekkerman, A., & Gilpin, G. (2011). Cost-Effective Hiring in U.S. High Schools: Estimating Optimal Teacher Quantity and Quality Decisions. SSRN Electronic Journal. doi:10.2139/ssrn.1873508

Bleed, J., & Wickline, M. R. (2006, August 22). Paron vote called attack on rural life. Arkansas Democrat-Gazette. Retrieved from http://www.highbeam.com/doc/1G1-149870947.html?refid=easy_hf

Chingos, M. M. (2012). Class Size and Student Outcomes: Research and Policy Implications. Journal of Policy Analysis and Management, 32(2), 411-438. doi:10.1002/pam.21677

Draper, N.R.; Smith, H. (1998). Applied Regression Analysis (3rd ed.). John Wiley. ISBN 0-471-17082-8.

Enders, C. K., & Tofighi, D. (2007). Centering Predictor Variables in Cross-Sectional Multilevel Models: A New Look at an Old Issue. Psychological Methods, 12(2), 121-138.

Fan, J., & Zhang, W. (2008). Statistical methods with varying coefficient models. Statistics and Its Interface, 1(1), 179-195.

Fodor, I. (2002) "A survey of dimension reduction techniques". Center for Applied Scientific Computing, Lawrence Livermore National, Technical Report UCRL-ID-148494

Fonti, V., & Belitser, E. (2017, March 30). Feature Selection using LASSO. VU Amsterdam. Retrieved from https://beta.vu.nl/nl/Images/werkstuk-fonti_tcm235-836234.pdf

Gold, M. R., Siegel, J. S., Russell, L. B., & Weinstein, M. C. (1996). Cost-Effectiveness in Health and Medicine. New York, NY: Oxford University Press.

Han, Kamber, Pei, Jaiwei, Micheline, Jian (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann.

Hanushek, E. A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. Educational Evaluation and Policy Analysis, 19(2), 141. doi:10.2307/1164207

Hendricks, M. D. (2013). Does it Pay to Pay Teachers More? Evidence from Texas. SSRN Electronic Journal. doi:10.2139/ssrn.2252576

Jha, U. K., Bajorski, P., Fokoue, E., Aardt, J. V., & Anderson, G. (2017). Dimensionality Reduction of High-Dimensional Highly Correlated Multivariate Grapevine Dataset. Open Journal of Statistics, 07(04), 702-717.

Krueger, A. (2002). Economic Considerations and Class Size. doi:10.3386/w887

Lake View School District v. Huckabee, 2002. No. 1992-5318

Nitta, K., Wrobel, S., & Holley, M. 2010. A Phenomenological Study of School Consolidation.

Office for Education Policy. (2010). Act 60: The Past, Present, and Future of School Consolidation in Arkansas. Policy Brief, 7(9).

Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. Journal of Public Economics, 130, 105-119. doi:10.1016/j.jpubeco.2015.02.008

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution, 1(2), 103-113.

Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Tow, C. (2006). The Effects of School Funding on Student Academic Achievement (K. Chay, Ed.) (Unpublished master's thesis). University of California, Berkeley.

United States Census Bureau; Community Facts. (2018). Retrieved February, 2018, from https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml

Wang, L., Li, H., & Huang, J. Z. (2008). Variable Selection in Nonparametric Varying-Coefficient Models for Analysis of Repeated Measurements. Journal of the American Statistical Association, 103(484), 1556-1569.

Wikipedia. "Efficient Frontier". Available at: https://en.wikipedia.org/wiki/Efficient_frontier. Retrieved October 16, 2017.

Zaki, M. J., & Meira, W., Jr. (2014). Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press.

Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. Annals of Translational Medicine, 4(7), 136-136.

7. Appendix

Table 6. All variables initially investigated with definition and source

Named again	Definition	Source
Average Commute Time	County mean travel time to work (minutes)	USGC
Below Poverty Line	Percentage of families and people whose income in the past 12 months is below the poverty level	USGC
Industry Type 1	Percent of civilians employed population 16 years and over that are in the industry: Agriculture, forestry, fishing and hunting, and mining	USGC
Industry Type 10	Percent of civilians employed population 16 years and over that are in the industry: Educational services, and health care and social assistance	USGC
Industry Type 11	Percent of civilians employed population 16 years and over that are in the industry: Arts, entertainment, and recreation, and accommodation and food services	USGC
Industry Type 12	Percent of civilians employed population 16 years and over that are in the industry: Other services, except public administration	USGC
Industry Type 13	Percent of civilians employed population 16 years and over that are in the industry: Public administration	USGC
Industry Type 2	Percent of civilians employed population 16 years and over that are in the industry: Construction	USGC
Industry Type 3	Percent of civilians employed population 16 years and over that are in the industry: Manufacturing	USGC
Industry Type 4	Percent of civilians employed population 16 years and over that are in the industry: Wholesale trade	USGC
Industry Type 5	Percent of civilians employed population 16 years and over that are in the industry: Retail trade	USGC
Industry Type 6	Percent of civilians employed population 16 years and over that are in the industry: Transportation and warehousing, and utilities	USGC
Industry Type 7	Percent of civilians employed population 16 years and over that are in the industry: Information	USGC
Industry Type 8	Percent of civilians employed population 16 years and over that are in the industry: Finance and insurance, and real estate and rental and leasing	USGC
Industry Type 9	Percent of civilians employed population 16 years and over that are in the industry: Professional, scientific, and management, and administrative and waste management services	USGC
County Occupation Cluster	cluster (0-2) for county based off of occupation percentages	USGC

Percent In Labor Force	Percent of population 16 years and over in labor force	USGC
Mean Income	County mean household income (dollars) for all households	USGC
Mean Income (Families)	County mean family income (dollars) for households with children under 18	USGC
Mean Income (With Earnings)	County mean income and benefits with earnings for all households	USGC
Median Income	County median household income (dollars) for all households	USGC
Median Income (Families)	County median family income (dollars) for households with children under 18	USGC
Occupation Type 1	Percent of civilians employed population 16 years and over with the occupation of: Management, business, science, and arts occupations	USGC
Occupation Type 2	Percent of civilians employed population 16 years and over with the occupation of: Service occupations	USGC
Occupation Type 3	Percent of civilians employed population 16 years and over with the occupation of: Sales and office occupations	USGC
Occupation Type 4	Percent of civilians employed population 16 years and over with the occupation of: Natural resources, construction, and maintenance occupations	USGC
Occupation Type 5	Percent of civilians employed population 16 years and over with the occupation of: Production, transportation, and material moving occupations	USGC
Parents In Labor Force	Percent of all families with children 6 to 17 years where all parents in family are in labor force	USGC
County Population	County population 16 years and over	USGC
Percent Unemployment	Percent of labor force 16 years and over that are unemployed	USGC
Cash Public Assistance	Percent of population who receive cash public assistance income	USGC
Food Stamps	Percent of population who receive Food Stamp/SNAP benefits in the past 12 months	USGC
Retirement Income	Percent of population who receive retirement income	USGC
Social Security	Percent of population who receive Social Security	USGC
Supplemental Security	Percent of population who receive Supplemental Security Income	USGC
Worker Class Type 1	Percent of civilians employed population 16 years and over that are Private wage and salary workers	USGC
Worker Class Type 2	Percent of civilians employed population 16 years and over that are Government workers	USGC
Worker Class Type 3	Percent of civilians employed population 16 years and over that are Self-employed in own not incorporated business workers	USGC

Worker Class Type 4	Percent of civilians employed population 16 years and over that are Unpaid family workers	USGC
Compensatory Education Expenditures	Expenditures for instructional activities designed primarily to meet the educational needs of pupils who are judged to be underachievers or educationally deprived. All compensatory education must be supplemental to regular instruction	ADE
Total Current Expenditures	Total Expenditures minus Capital Expenditures minus Debt Service	ADE
Extracurricular Expenditures	Expenditures for extracurricular activities	ADE
Facilities Expenditures	Expenditures for activities concerned with acquiring land and buildings, remodeling buildings, constructing buildings and additions to buildings, initially installing or extending service systems, and site improvements	ADE
Total (Calculated) Expenditures	Total Expenditures + Total District Level Support + Total School Level Support + Total Non-Instructional Services + Facilities Acquisition and Construction + Debt Service + Other Non-Programmed Costs	ADE
Total (From ADE) Expenditures	Net current expenditures divided by the four-quarter Average Daily Attendance (ADA). Arkansas uses the three-quarter Average Daily Membership (ADM) for funding and other analytical purposes. Users of this information should be aware of this difference	ADE
State Foundation Revenue	Per-student amount of state financial aid provided to a school district under § 6-20-2305(a)(1)	ADE
Isolated Revenue	State financial aid provided to isolated school districts, small school districts, or districts with isolated school areas as set forth in A.C.A. §§ 6-20-601 et seq. and restricted for use by those isolated school districts, small school districts, or districts with isolated school areas	ADE
Other Revenue	Financing Sources + Balances from Consolidation/Annexed District + Indirect Cost Reimbursement + Gains and Losses from Sale of Fixed Assets + Compensation for Loss of Fixed Assets + Other	ADE
Property Tax Revenue	Revenue comprised of property taxes, property tax relief, tax accruals, delinquent taxes, excess commissions, land redemptions, penalties and interest on delinquent taxes, and other local taxes	ADE
Federal Restricted Revenue	Restricted funds provided by the federal government through the state as agent to the school districts, which must be used for specific categorical purposes, such as revenue in lieu of taxes, Elementary/Secondary Education Programs, ROTC, Carl Perkins Stabilization Aid, Adult Education Stabilization, School Food Services, IDEA Title VI, and Safe and Drug Free Schools	ADE

State Restricted Revenue	Adult Education plus Professional Development + Other Regular Education + Gifted and Talented + Alternative Learning Environment + English Language Learners + National School Lunch Categorical + Other Special Education + Career Education + School Food Service + Education Service Cooperatives + Early Childhood Programs + Magnet School Programs + Other Non-instructional Program Aid	ADE
Initial Revenue	Total revenue divided by the number of students	ADE
Total Revenue	Total Unrestricted Revenue + Total Restricted Revenue from State Sources + Total Restricted Revenue from Federal Sources + Total Other Sources of Income	ADE
Unrestricted Revenue	The total revenue of state unrestricted funds	ADE
Advance Degree	Percent of teachers that have an advanced degree	ADE
Masters or Advance Degree	Percent of teachers that have an master's degree or an advanced degree	ADE
Salary Increase	Average teacher salary increase per year calculated from the district teacher salary schedule	ADE
Bachelors Degree	Percent of teachers that have an bachelor's degree	ADE
Classified Staff Ratio	Classified staff total divided by number of students	ADE
Classified Staff	Total number of classified staff; any employee who performs work for the school district under a written annual contract in a position that does not require a valid teaching license issued by the Arkansas State Board	ADE
Completely Certified	Percent of teachers that are completely certified as defined by ADE	ADE
Provincial Credentials	Percent of teachers that are teaching using an emergency or provisional credential	ADE
Licensed Staff Ratio	Total number of licensed staff; a person hired by the local school district who is compelled by law or regulation to secure a license from the State Board of Education.	ADE
Masters Degree	Percent of teachers that have an master's degree	ADE
Average Salary	Average teacher salary	ADE
Minimum Salary	Salary of a teacher with a bachelor's degree and no years of experience. This is the minimum salary.	ADE
Total Teachers	Total teachers at a school	ADE
Unqualified Teachers	Percent of teachers that are highly qualified as defined by ADE	ADE
Average Years Experience	Average years of teacher experience	ADE
Accreditation Status	Accreditation school status	ADE
Alternative Status	Alternative school status	ADE
Block Schedule Status	Block schedule school status	ADE
Normalized Test Score	Normalized test score from combined math and literacy scores	ADE

Federal Program Status	Federal program school status	ADE
Isolated Status	Binary variable; equal to 1 if the district receives isolated funding and 0 otherwise	ADE
LEA	Unique identifier for each school	ADE
Letter Grade	School letter grade given by ADE based on school performance	ADE
Letter Grade Points	Points calculated for school letter grade given by ADE based on school performance	ADE
Magnet Status	Magnet school status	ADE
Night Status	Night school status	ADE
Student Teacher Ratio (From ADE)	Student teacher ratio as calculated by ADE (rounded down)	ADE
Year Round Status	Year round school status	ADE
American Indian Percent	Percent of students with the following race: An American Indian or Alaska Native person has origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment. (NCES.ed.gov)	ADE
Asian Percent	Percent of students with the following race: An Asian person has origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent, including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam. (NCES.ed.gov)	ADE
Disciplinary Actions	Total disciplinary actions recorded over the school year	ADE
Female Percent	Percent of students with the following attribute: Gender selection of female	ADE
G/T and Free/Reduced Percent	Percent of students in the gifted and talented program who receive free or reduced lunches	ADE
Grade 1	Number of students in 1st Grade enrolled	ADE
Grade 10	Number of students in 10th Grade enrolled	ADE
Grade 11	Number of students in 11th Grade enrolled	ADE
Grade 12	Number of students in 12th Grade enrolled	ADE
Grade 2	Number of students in 2nd Grade enrolled	ADE
Grade 3	Number of students in 3rd Grade enrolled	ADE
Grade 4	Number of students in 4th Grade enrolled	ADE
Grade 5	Number of students in 5th Grade enrolled	ADE
Grade 6	Number of students in 6th Grade enrolled	ADE
Grade 7	Number of students in 7th Grade enrolled	ADE
Grade 8	Number of students in 8th Grade enrolled	ADE
Grade 9	Number of students in 9th Grade enrolled	ADE
Grade None	Number of students in not enrolled in a grade	ADE
Married Percent	Percent of students with the following attribute: Legally married	ADE

Orphan Percent	Percent of students with the following attribute: Denotes a student with no living paternal parents	ADE
Pacific Islander Percent	Percent of students with the following race: An Other Pacific Islander or Native Hawaiian person has origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands. (NCES.ed.gov)	ADE
Single Percent	Percent of students with the following attribute: A student who is not legally married	ADE
Student Teacher Ratio (Calculated)	Actual student teacher ratio calculated as number of teachers divided by total enrollment at a school	ADE
Two Races Percent	Percent of students with the following race: Two or more races were selected	ADE
Attendance Rate	Attendance rate of students at the school	ADE
Bilingual Percent	% Students Who Speak a Language Other Than English	ADE
Black Percent	Percent of students with the following race: An African American or Black person has origins in any of the black racial groups of Africa. (NCES.ed.gov)	ADE
Disciplinary Actions Ratio	Total disciplinary actions recorded divided by school total enrollment	ADE
Enrollment Change	Enrollment % Change from Oct 1 to Final of a school	ADE
Foster Percent	Percent of students with the following attribute: Refers to a student that lives in a foster home environment	ADE
Free/Reduced Lunch Percent	Percent of students who receive free or reduced lunches	ADE
Gifted/Talented Percent	Percent of students in the gifted and talented program who receive free or reduced lunches	ADE
Handicapped Percent	Percent of students with the following attribute: A student has been determined to be eligible under Section 504 of the Rehabilitation Act of 1973. For purposes of this database this does not include special education students	ADE
Hispanic Percent	Percent of students with the following race: A Hispanic or Latino person is of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race. (NCES.ed.gov)	ADE
Homeless Percent	Percent of students with the following attribute: Homeless	ADE
Male Percent	Percent of students with the following attribute: Gender selection of male	ADE
Migrant Percent	Percent of students with the following attribute: A student who has moved in the past 3 years, on their own or with their family, for the purpose of seeking work in agriculture, fishing, dairies, logging or food processing. A student can only be determined as "migrant" by the Arkansas migrant education program, which will provide a list of eligible students to each district where migrant children reside	ADE

Military Family Percent	Percent of students with the following attribute: Parents are in the military	ADE
English Learners Percent	Percent of students with the following attribute: The student has a language background other than English, and his or her proficiency in English is such that the probability of the student's academic success in an English-only classroom is below that of native English language students	ADE
Other Race Percent	Percent of students with the following race: A student selected a race other than Black/African American, Hispanic/Latino, or White	ADE
Pre-K Enrollment	Number of students who are enrolled in a pre-kindergarten program at the school	ADE
School Choice Percent	Percent of students at the school who attend the school out of their assigned district	ADE
Special Education Percent	Percent of students with the following attribute: Students who receive special education services	ADE
Total Students	Total students attending the school	ADE
White Percent	Percent of students with the following race: A White person has origins in any of the original peoples of Europe, the Middle East, or North Africa. (NCES.ed.gov)	ADE
School Year	School year	ADE

Figure 1-A. Results from Box-Cox transformation performed on 2013 test scores

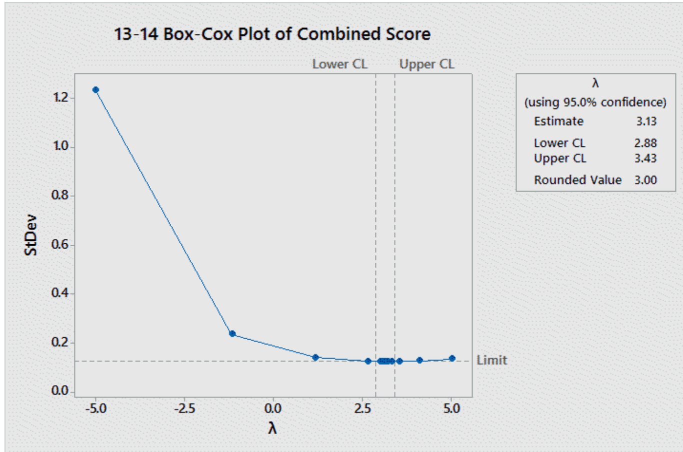


Figure 1-B. Results from Box-Cox transformation performed on 2014 test scores

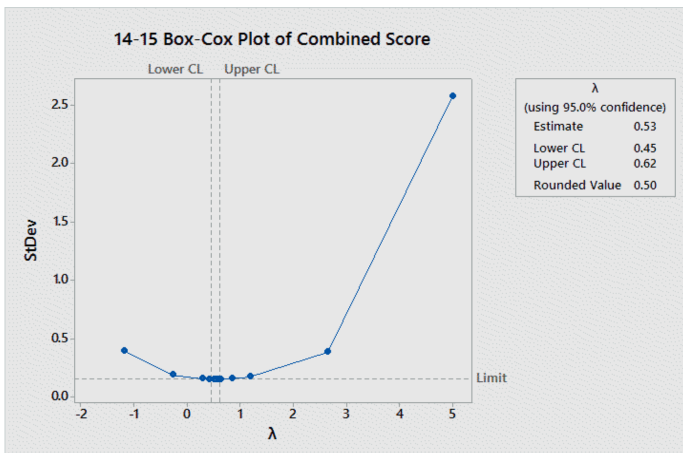
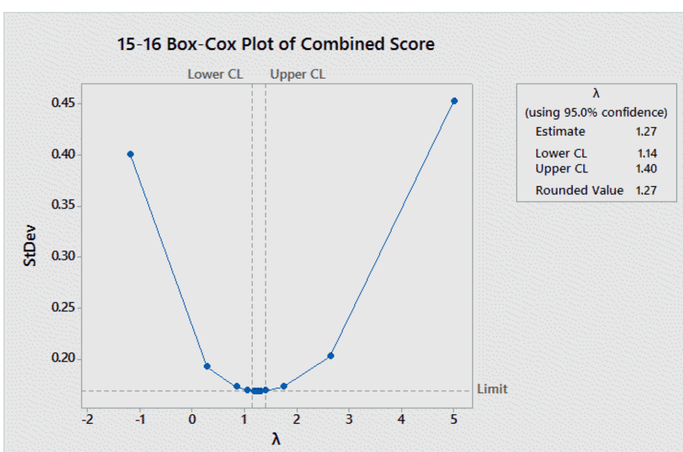


Figure 1-C. Results from Box-Cox transformation performed on 2015 test scores



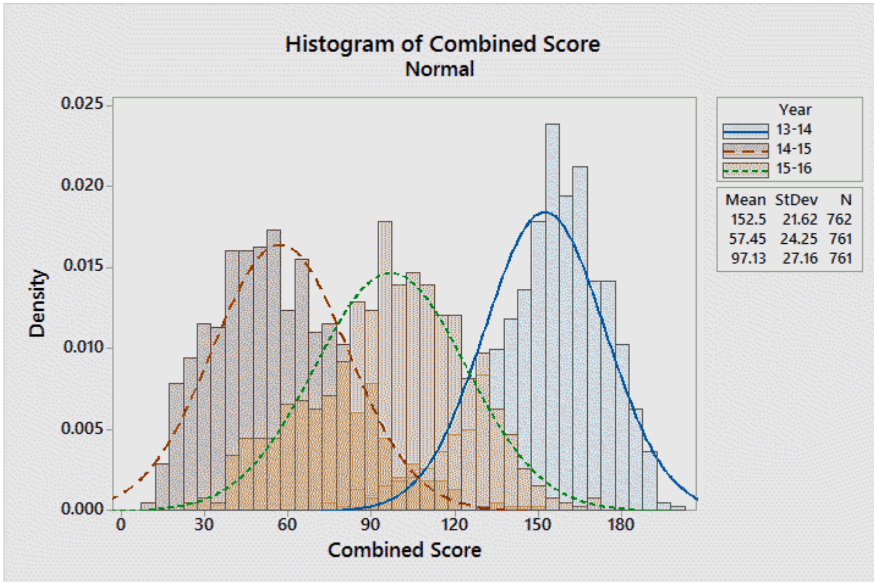


Figure 2-A. Test score distribution before Box-Cox transformation

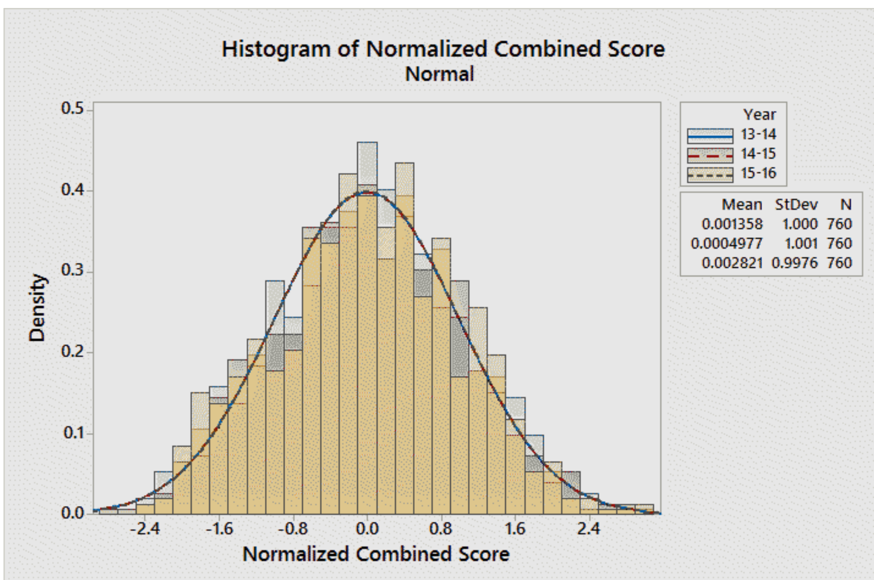


Figure 2-B. Test score distribution after Box-Cox transformation and normalization

Table 7. Results from k-means clustering on county occupation statistics

County Cluster	Count
0	309
1	702
2	1269

Figure 3-A. Inertia for different values of k from clustering on county occupation statistics

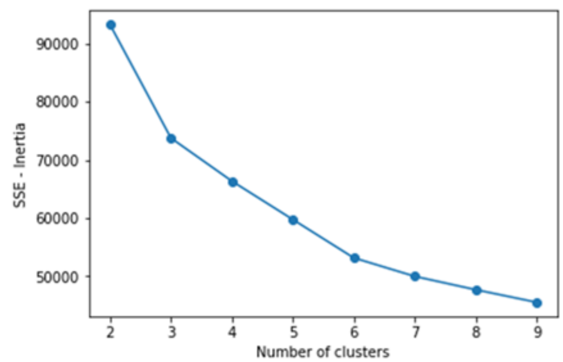


Figure 3-B. Calinski-Harabaz score for different values of k from clustering on county occupation statistics

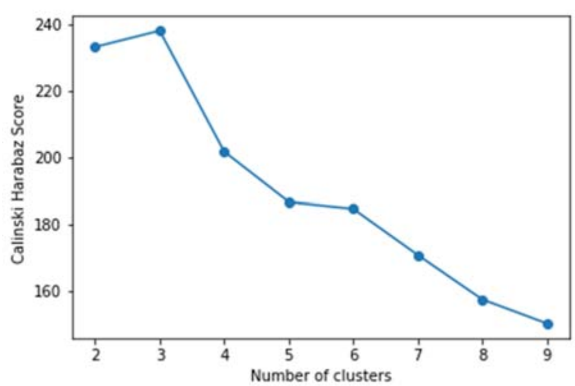


Figure 3-C. Silhouette score for different values of k from clustering on county occupation statistics

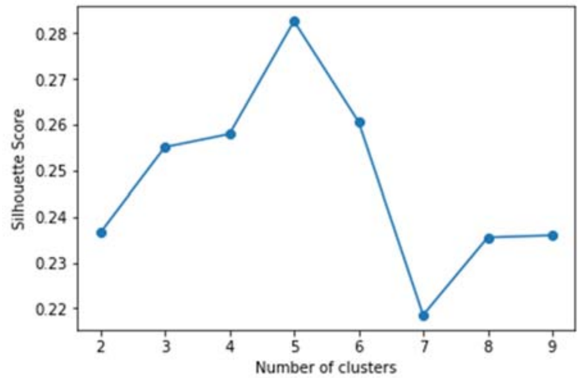


Table 8. Significant interaction terms

Variable 1	Variable 2
county_job_cluster	county_population
county_job_cluster	students_attendence_rate
county_job_cluster	students_black_percent
county_job_cluster	students_hispanic_percent
county_job_cluster	students_free_reduced_lunch_percent
school_isolated_funding_binary	county_w_food_stamps_percent
school_isolated_funding_binary	finance_revenue_propertytax
school_isolated_funding_binary	students_specialed_percent
county_avg_commute_time_percent	finance_expense_compensatory
county_avg_commute_time_percent	finance_revenue_restricted_state
county_avg_commute_time_percent	students_black_percent
county_median_income_families_estimate	finance_expense_facilities
county_median_income_families_estimate	personnel_salary_avg
county_median_income_families_estimate	personnel_advance_degree_percent
finance_revenue_propertytax	students_other_percent
finance_expense_compensatory	students_hispanic_percent
finance_expense_facilities	students_specialed_percent
finance_expense_facilities	personnel_classified_staff_percent
finance_revenue_restricted_state	students_other_percent
finance_revenue_restricted_state	students_hispanic_percent
finance_revenue_unrestricted	students_other_percent
personnel_salary_avg	students_attendence_rate
personnel_salary_avg	students_free_reduced_lunch_percent
student_teacher_ratio	students_free_reduced_lunch_percent
students_other_percent	students_black_percent
students_other_percent	students_hispanic_percent
students_other_percent	students_free_reduced_lunch_percent
students_specialed_percent	students_disciplinary_actions_percent
students_free_reduced_lunch_percent	students_free_reduced_lunch_percent

Table 9. Linear regression coefficients and significance for each model

Variable	2013-2014			2014-2015			2015-2016		
	Coef	SE Coef	P-Value	Coef	SE Coef	P-Value	Coef	SE Coef	P-Value
constant	0.0014	0.0221	0.951	0.0005	0.0234	0.983	0.0028	0.0209	0.893
County Occupation Cluster	0.274	0.122	0.025	0.062	0.128	0.629	0.1	0.118	0.399
Isolated Status	0.0085	0.0284	0.766	0.0025	0.0333	0.941	0.0145	0.0268	0.589
County Population	-0.667	0.26	0.011	-0.139	0.275	0.614	-0.104	0.256	0.686
Average Commute Time	-0.0722	0.0367	0.05	-0.065	0.0423	0.125	-0.0447	0.0359	0.213
Food Stamps	0.0986	0.0509	0.053	0.065	0.0538	0.227	0.0654	0.0487	0.179
Median Income (Families)	0.1646	0.0641	0.01	0.1324	0.0693	0.056	-0.0185	0.0629	0.769
Property Tax Revenue	0.1135	0.0596	0.057	0.369	0.0785	0	0.3752	0.0642	0
Compensatory Education Expenditures	0.0374	0.0347	0.281	-0.0014	0.0366	0.969	-0.015	0.0331	0.65
Facilities Expenditures	0.0043	0.0294	0.884	-0.0372	0.0372	0.318	0.0355	0.0323	0.271
State Restricted Revenue	-0.0167	0.0411	0.685	0.1309	0.0501	0.009	0.0144	0.0438	0.742
Unrestricted Revenue	-0.0778	0.0592	0.189	-0.2605	0.0757	0.001	-0.242	0.0601	0
Average Salary	0.2222	0.0485	0	0.1942	0.0538	0	0.1149	0.0482	0.017
Student Teacher Ratio (Calculated)	0.2525	0.033	0	0.1282	0.0347	0	0.1023	0.0315	0.001
Attendance Rate	0.0167	0.0298	0.575	-0.0032	0.0331	0.923	0.0049	0.0268	0.856
Other Race Percent	0.0629	0.0379	0.098	0.0359	0.0403	0.373	0.0024	0.0383	0.949
Black Percent	-0.4753	0.0596	0	-0.3583	0.0624	0	-0.4205	0.0553	0
Hispanic Percent	-0.079	0.0391	0.044	-0.0388	0.0423	0.359	-0.0076	0.0374	0.84
Foster Percent	-0.0207	0.0241	0.39	-0.0022	0.0253	0.93	-0.0369	0.0235	0.116
Male Percent	-0.0225	0.0231	0.329	-0.0583	0.0248	0.019	-0.0337	0.0218	0.124
Special Education Percent	-0.0913	0.0276	0.001	-0.0394	0.0295	0.182	-0.0808	0.026	0.002
Total Students	-0.1187	0.0323	0	-0.1051	0.035	0.003	-0.0563	0.0304	0.065
Disciplinary Actions Ratio	-0.141	0.0265	0	-0.0809	0.0291	0.006	-0.1025	0.028	0
Free/Reduced Lunch Percent	-0.2205	0.0526	0	-0.3235	0.0564	0	-0.3194	0.0487	0
Classified Staff Ratio	0.1434	0.0281	0	0.0794	0.0291	0.007	0.0104	0.0269	0.698
Advance Degree	-0.0227	0.0277	0.413	-0.0424	0.0291	0.145	0.0359	0.0249	0.15
Average Years Experience	0.0061	0.0266	0.818	0.0285	0.0292	0.328	0.0669	0.0268	0.013
Completely Certified	0.0668	0.0252	0.008	0.0085	0.0266	0.75	-0.0145	0.0239	0.544
County Occupation Cluster * County Population	0.35	0.209	0.094	-0.071	0.222	0.749	0.02	0.207	0.924
County Occupation Cluster * Attendance Rate	0.0269	0.0291	0.356	0.0539	0.0325	0.097	0.0067	0.0264	0.8
County Occupation Cluster * Black Percent	-0.0064	0.0383	0.868	-0.052	0.0426	0.222	-0.0801	0.0368	0.03
County Occupation Cluster * Hispanic Percent	-0.152	0.0378	0	-0.1222	0.0416	0.003	-0.1059	0.036	0.003
County Occupation Cluster * Free/Reduced Lunch Percent	0.0274	0.0534	0.607	-0.0066	0.0572	0.909	0.0144	0.0491	0.77
Isolated Status * Food Stamps	0.0244	0.0276	0.377	0.0699	0.0291	0.017	0.0379	0.0242	0.117
Isolated Status * Property Tax Revenue	-0.0152	0.0309	0.622	-0.0824	0.0316	0.009	-0.053	0.0265	0.046
Isolated Status * Special Education Percent	0.0137	0.026	0.599	-0.0822	0.027	0.002	-0.0486	0.0239	0.042

Average Commute Time * Compensatory Education Expenditures	0.0203	0.0308	0.509	0.0785	0.0325	0.016	0.0374	0.0308	0.224
Average Commute Time * State Restricted Revenue	0.0343	0.0296	0.247	0.0288	0.0341	0.398	-0.0123	0.0291	0.674
Average Commute Time * Black Percent	0.0336	0.0364	0.356	-0.0073	0.0357	0.838	0.0217	0.0302	0.472
Median Income (Families) * Facilities Expenditures	-0.0278	0.0299	0.352	-0.1013	0.0418	0.016	-0.0631	0.0252	0.013
Median Income (Families) * Average Salary	0.0761	0.0325	0.02	0.0622	0.0365	0.088	0.0566	0.0338	0.095
Median Income (Families) * Advance Degree	0.0174	0.0272	0.524	0.039	0.0285	0.172	0.0084	0.0248	0.735
Property Tax Revenue * Other Race Percent	0.0584	0.0551	0.29	0.0548	0.0736	0.456	0.0723	0.0612	0.237
Compensatory Education Expenditures * Hispanic Percent	-0.104	0.0291	0	-0.0542	0.0309	0.08	-0.0336	0.0278	0.227
Facilities Expenditures * Special Education Percent	0.0548	0.026	0.035	0.0167	0.0277	0.546	0.0068	0.0249	0.786
Facilities Expenditures * Classified Staff Ratio	0.0168	0.0337	0.617	0.0953	0.0358	0.008	0.0726	0.029	0.012
State Restricted Revenue * Other Race Percent	0.0773	0.0301	0.01	0.0208	0.0341	0.542	0.0711	0.0295	0.016
State Restricted Revenue * Hispanic Percent	-0.0412	0.0261	0.114	-0.049	0.029	0.091	-0.0611	0.0265	0.021
Unrestricted Revenue * Other Race Percent	-0.0276	0.057	0.628	0.0546	0.0733	0.457	-0.0189	0.0606	0.756
Average Salary * Attendance Rate	-0.0061	0.0296	0.837	0.0549	0.0289	0.057	0.0546	0.0245	0.026
Average Salary * Free/Reduced Lunch Percent	0.0612	0.0398	0.124	0.0067	0.0448	0.882	-0.0163	0.0436	0.709
Student Teacher Ratio (Calculated) * Free/Reduced Lunch Percent	-0.0127	0.0302	0.674	-0.0267	0.0304	0.38	-0.0278	0.0279	0.32
Other Race Percent * Black Percent	0.0178	0.0456	0.697	0.0596	0.0495	0.229	-0.005	0.0436	0.908
Other Race Percent * Hispanic Percent	-0.0894	0.0407	0.028	-0.0341	0.0471	0.469	-0.0662	0.0451	0.142
Other Race Percent * Free/Reduced Lunch Percent	-0.0065	0.0363	0.858	-0.0122	0.041	0.765	0.0251	0.0379	0.508
Special Education Percent * Disciplinary Actions Ratio	0.0446	0.024	0.063	0.0291	0.0252	0.248	0.0376	0.0232	0.105
Free/Reduced Lunch Percent * Free/Reduced Lunch Percent	0.0844	0.031	0.007	0.0243	0.0347	0.485	0.0844	0.0308	0.006

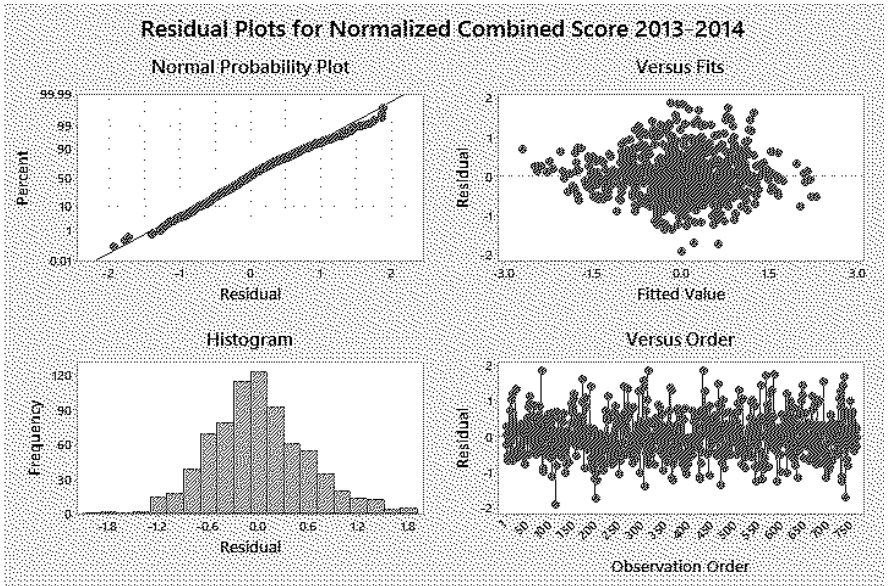


Figure 4-A. Residual plots for the multiple linear regression on 2013-2014 school year data

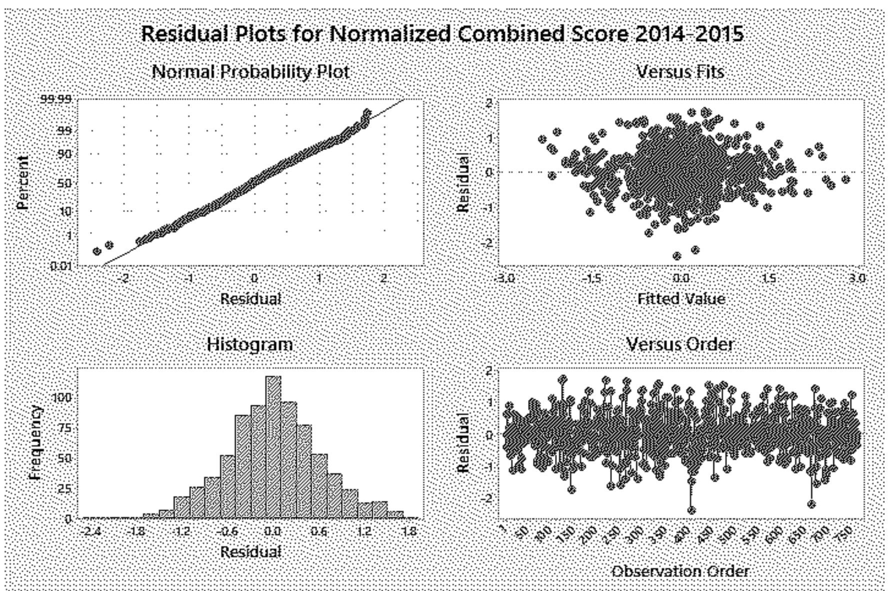


Figure 4-B. Residual plots for the multiple linear regression on 2014-2015 school year data

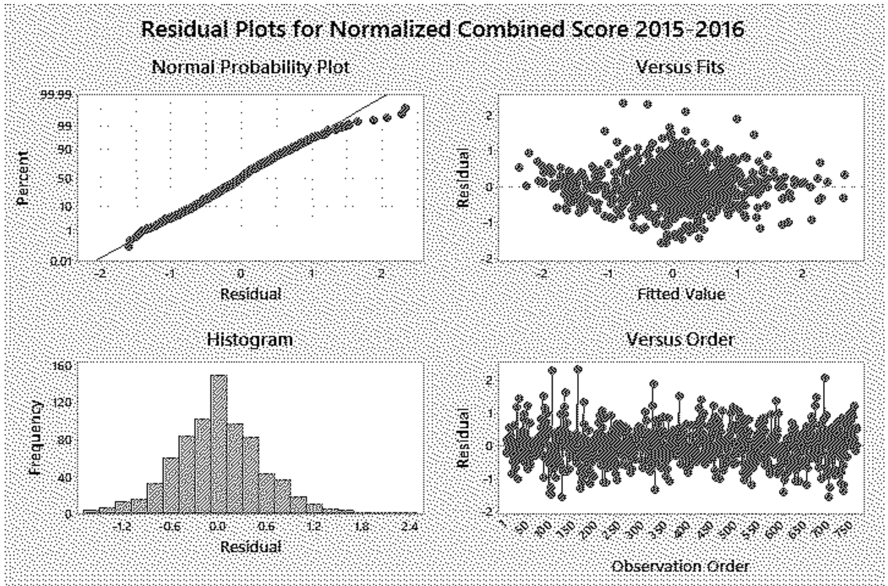


Figure 4-C. Residual plots for the multiple linear regression on 2015-2016 school year data