A Mathematical Model for Substitute Teacher Demands

Daniel J. Littleton

University of North Texas
Texas Eta Chapter

Aletheia—The Alpha Chi Journal of Undergraduate Scholarship

- This publication is an online, peer-reviewed, interdisciplinary undergraduate journal, whose mission is to promote high quality research and scholarship among undergraduates by showcasing exemplary work.
- Submissions can be in any basic or applied field of study, including the physical and life sciences, the social sciences, the humanities, education, engineering, and the arts.
- Publication in *Aletheia* will recognize students who excel academically and foster mentor/mentee relationships between faculty and students.
- In keeping with the strong tradition of student involvement in all levels of Alpha Chi, the journal will also provide a forum for students to become actively involved in the writing, peer review, and publication process.
- More information can be found at www.alphachihonor.org/aletheia. Questions to the editors may be directed to tlindblom@alphachihonor.org or kvosevich@alphachihonor.org.

Alpha Chi is a national college honor society that admits students from all academic disciplines, with membership limited to the top 10 percent of an institution's juniors, seniors, and graduate students. Invitation to membership comes only through an institutional chapter. A college seeking a chapter must grant baccalaureate degrees and be regionally accredited. Some 300 chapters, located in almost every state, induct approximately 11,000 members annually. Alpha Chi members have been "making scholarship effective for good" since 1922.

ALETHEIA
The Alpha Chi
Journal of
Undergraduate
Scholarship

A Mathematical Model for Substitute Teacher Demands

Daniel J. Littleton

University of North Texas
Texas Eta Chapter

Abstract

The following mathematical model provides a method to determine the most cost efficient means of providing human resources for a designated school district. The author utilizes a mathematical model to find the most efficient solution to a designated resource problem (the need for substitute teachers). The author then generates a graphical representation of the raw data, which he then interprets and presents to the project supervisor. The completion of this model is the result of the collection and analysis of human resource data from the Denton Independent School District and the execution of a computer script configured specifically for this project. The final goal is the determination of the optimal pool size of substitute teachers that can fill positions while not resulting in excess cost to the district. The final results of the model demonstrate that the pool size is dependent upon the pay rates of the substitutes hired to perform their duties, as well as the pay rates of full-time teachers called upon to fill vacant posts should the pool size prove to be inadequate to the district's needs.

Keywords: Keywords: Model, Simulated Demands, Inverse Linear Splines, Monte Carlo Simulation

Model Parameters

The purpose of this analysis is to find the optimal population size for a pool of substitute teachers within Denton Independent School District (DISD). Substitute teachers conduct classroom instruction when the regular teacher is not present. The substitutes are compensated for each full day of work that they log. Occasionally, full-time teachers are required to supervise additional classes when the substitute pool is insufficiently staffed.

According to the Employee Handbook for DISD, Professional Teachers do not normally qualify for overtime pay when they are required to provide instruction for extra classes. However, for the purposes of this model, it shall be assumed that there are special circumstances in which a Regular Full Time Teacher may qualify for overtime pay. According to the Handbook, overtime pay is considered to be a per diem supplement to the daily pay rate for the individual teacher. This supplement is a fixed rate of 20 dollars for each class that the full-time teacher is told to instruct in addition to their regular schedule. It shall be assumed that each full-time teacher shall not be required to cover more than a single extra class in a single day in addition to their regular schedule.

In order to minimize the costs to the school district, it is optimal to have enough substitute teachers in the pool to provide coverage for all absent teachers during the school year. This will eliminate the need for the district to provide overtime pay to their regular teachers. However, this is most often not the case for any school district.

Modeling Procedure

This model uses data collected from the school district to find the optimal substitute pool size that minimizes the costs of providing instructors for all classes. The data is evaluated using Simulated Demands, and the simulations are defined using Random Numbers, Linear Splines, and Inverse Linear Splines (Fox, Giordano, & Horton, 2014). The simulations are separated into 10 evenly-spaced intervals using Cumulative Percentage. Once the relative and cumulative percentages are established, the Monte Carlo Simulation (Fox, Giordano, & Horton, 2014) is employed to determine the optimal substitute pool size to minimize costs.

Each simulation requires a unique range, interval, and number of evaluations in order to provide an accurate graphical representation of the data. The range of the simulation refers to the separate closed intervals of the substitute teacher population. The intervals are derived from the raw data provided by DISD, as are the endpoints of each interval. As these are closed intervals, there exists a minimum and maximum point for each data set. The minimum population of each interval is the left endpoint of the interval, and the maximum population of each interval is the right endpoint. The interval width utilized by the model is adjusted for each evaluation of the script to create a precise graph of the Inverse Linear Spline. This adjustment is also implemented to expedite the compiling of the script.

The number of evaluations refers to the number of instances that the script will repeat the Monte Carlo Simulation to determine an accurate Inverse Linear Spline function. This function is the expression utilized to create a precise graphical representation of the substitute teacher pool population and the total cost to the district, as well as how these values change as the population of the pool of substitutes and the pay rates of the relevant employees vary. Therefore, each simulation shall be provided with a description of all parameters used to compile the script associated with it. The final goal of the model is to determine the optimal pool size and the corresponding cost of application for DISD.

The graphical representation of the data is produced by a MATLAB code that processes the data inputs and compiles the script of the Monte Carlo Simulation. MATLAB is a computer application that allows the user to plot mathematical models in the form of a chart or graph. The model is based upon a specified interval, and the associated cost in dollars for each data point within the interval. The data points are indicated on the graph of the Inverse Linear Spline. Each point specifies a population size for the substitute teacher pool and the associated cost to DISD to maintain said pool. The optimal number of substitutes is defined as the point that graphs to the lowest point of the functions' trough. This point is a population value defined by the function, while maintaining minimal cost to the district.

The pay rates for all the relevant teachers employed by the district are provided in the following table:

Employee Title	Daily Pay Rate
Certified Substitute Teacher	90
Degreed Substitute Teacher	80
Non-Degreed Substitute Teacher	70
Teacher with One Year Experience Overtime	299.41
Teacher with 10 Years' Experience Overtime	314.25

(Denton Independent School District Human Resources, 2017)

The number of substitutes needed by the district to fill vacancies are tracked on a weekly basis during each academic year. These figures were provided by the Human Resources Department of DISD. The logged figures of substitute requirements by the district are provided in the following table:

Weekly counts from 08/21/16 thru 06/03/17	Needed Substitutes
8/21/2016 - 8/27/2016	479
8/28/2016 - 9/3/2016	677
9/4/2016 - 9/10/2016	607
9/11/2016 - 9/17/2016	909
9/25/2016 - 10/1/2016	892
10/2/2016 - 10/8/2016	841
10/9/2016 - 10/15/2016	891
10/16/2016 - 10/22/2016	1192
10/23/2016 - 10/29/2016	1161
10/30/2016 - 11/5/2016	1093
11/6/2016 - 11/12/2016	1013
11/13/2016 - 11/19/2016	1070
11/27/2016 - 12/3/2016	1013
12/4/2016 - 12/10/2016	974
12/11/2016 - 12/17/2016	704
1/1/2017 - 1/7/2017	679
1/8/2017 - 1/14/2017	1063
1/15/2017 - 1/21/2017	764
1/22/2017 - 1/28/2017	1101
1/29/2017 - 2/4/2017	1072
2/5/2017 - 2/11/2017	1344
2/12/2017 - 2/18/2017	1120
2/19/2017 - 2/25/2017	1262
2/26/2017 - 3/4/2017	1167
3/5/2017 - 3/11/2017	1141
3/19/2017 - 3/25/2017	950
3/26/2017 - 4/1/2017	972
4/2/2017 - 4/8/2017	1272
4/9/2017 - 4/15/2017	862
4/16/2017 - 4/22/2017	843
4/23/2017 - 4/29/2017	1083
4/30/2017 - 5/6/2017	1104
5/7/2017 - 5/13/2017	1159
5/14/2017 - 5/20/2017	1487
5/21/2017 - 5/27/2017	1187
5/28/2017 - 6/3/2017	311

Cumulative Percentage Demand Intervals

The collected data can be interpreted in a manner that allows for a Monte Carlo Simulation. By analyzing the data trends, a relative and cumulative percentage can be determined which allows for a more efficient modeling structure of the data trends. The relative percentages were determined using the basic principles of percent theory, where the number of data points in the designated interval is divided by the total number of data points provided. First, an interval for the data must be selected which shall be used to extrapolate a Relative Percentage. Secondly, the Cumulative Percentage is determined by the summation of each Relative Percentage as it appears in the sequence. Upon review of the data, an interval of 117 was chosen to establish the following percentage parameters:

Interval (Increments of 117)	Relative Percentage	Cumulative Percentage
311 – 428	2.7	2.7
429 – 546	2.7	5.4
547 – 664	2.7	8.1
665 – 782	10.8	18.9
783 – 900	13.5	32.4
901 – 1018	16.2	48.6
1019 – 1136	24.4	73
1137 – 1254	16.2	89.2
1255 – 1372	8.1	97.3
1373 – 1490	2.7	100

Modeling the Optimal Pool Size for Substitute Teachers

A model for the optimal pool size of Substitute Teachers is created from the established parameters utilizing a MATLAB script. The script processes the Intervals, Relative and Cumulative Percentages, and Pay Rates of Full Time and Substitute Teachers as variables to produce data points. The MATLAB script created by the author is listed below. This script was employed specifically for the first iteration of Non-Degreed Substitutes and One Year Teachers. The notations next to each line of the script, preceded by a % symbol, are comments used to guide the understanding of both the author and the reader.

```
%End Points of the Intervals
h = 117; %Interval Width
a = 311:h:1373; %Left endpoint of each interval
b = 428:h:1490; %Right endpoint of each interval
%Relative and Cumulative Percentages
rp = [2.7 2.7 2.7 10.8 13.5 16.2 24.4 16.2 8.1 2.7]'; %Relative Percentage
cp = [2.7 5.4 8.1 18.9 32.4 48.6 73 89.2 97.3 100]/100'; %Cumulative Percentage
n = length(cp); %The number of intervals evaluated by the Monte Carlo Simulation
%Determine the linear splines and their inverse:
%The first Inverse Linear Spline covers half an interval, the last covers 1.5 intervals
%The kth piece of the Inverse Linear Spline is given by the following function
%y = x0 + (rd-y0)/m where m is the slope of the spline,
```

%rd is a random number, (x0,y0) is the left midpoint of an individual Linear Spline interval
%Calculate slopes
m(1) = cp(1)/(h/2); %First Interval
m(n) = (cp(n) - cp(n-1))/(1.5*h); %Last Interval
for k = 2:n-1
    m(k) = (cp(k) - cp(k-1))/h;
end
%Run the simulation for the different pool size S for N times
%Use the average cost as the cost for that pool size

N = 1000;
%Shifting the cost shifts the whole index and results in a new model
Cost = [];
p = 70; %Pay Rate for substitutes
r = 299.41; %Pay rate for overtime teachers
%NOTE: AS THE DATA IS MODELED A MORE ACCURATE INTERVAL OF S CAN BE
%DETERMINED BASED UPON THE RESULTING PLOTS. IN ORDER S IS 300:100:1500,
%900:25:1200, 1050:5:1125, 1095:1:1105
Sk = 1;
for S = 300:100:1500 %S shifts according to the precision given by the data
    C = 0; %Cost for each S
    for k = 1:N %This for loop is the beginning of the Monte Carlo Simulation
        rd = rand; %Each random number will produce a demand X
        if rd < cp(1), x0 = (a(1) + b(1))/2; y0 = cp(1); x = x0 + (rd-y0)/m(1);
        elseif rd < cp(2), x0 = (a(2) + b(2))/2; y0 = cp(2); x = x0 + (rd-y0)/m(2);
        elseif rd < cp(3), x0 = (a(3) + b(3))/2; y0 = cp(3); x = x0 + (rd-y0)/m(3);
        elseif rd < cp(4), x0 = (a(4) + b(4))/2; y0 = cp(4); x = x0 + (rd-y0)/m(4);
        elseif rd < cp(5), x0 = (a(5) + b(5))/2; y0 = cp(5); x = x0 + (rd-y0)/m(5);
        elseif rd < cp(6), x0 = (a(6) + b(6))/2; y0 = cp(6); x = x0 + (rd-y0)/m(6);
        elseif rd < cp(7), x0 = (a(7) + b(7))/2; y0 = cp(7); x = x0 + (rd-y0)/m(7);
        elseif rd < cp(8), x0 = (a(8) + b(8))/2; y0 = cp(8); x = x0 + (rd-y0)/m(8);
        elseif rd < cp(9), x0 = (a(9) + b(9))/2; y0 = cp(9); x = x0 + (rd-y0)/m(9);
        else x0 = b(10); y0 = cp(10); x = x0 + (rd-y0)/m(10);
        end % End of the initial if statement
        if x < S, Cxs = p*S; else Cxs = p*S + (x-S)*r; end %Cost for the modeled demand
        C = C + Cxs; %Summation of the cost for each simulation
    end %End of the k interval

    Cost(Sk) = C/N; %Uses the average cost as the cost for that point in S
    Sk = Sk +1;
end %End of the for loop and the Monte Carlo Simulation
plot(Cost,'*') %Plot the data points for the Population of Substitutes vs. The Cost to DISD
hold on %Hold the plot of the Cost Data Points
plot(Cost) %Plot the curve of the Inverse Linear Spline intervals found via Monte Carlo Simulation

Once the variables have been processed by MATLAB (in order to model the data points on a specified interval), the points are then critiqued to narrow the interval upon which the pool size is determined. The intervals are presented in numeric form and separated by a colon between each value. To better understand

the application of these intervals, the interval listed as 300:100:1500 may be interpreted in the following manner: 300 is the starting point for the population of the substitute teacher pool while 1500 is the ending point. In essence, the MATLAB script will process the optimal pool of substitutes as existing between 300 and 1500 substitutes. The 100 listed between the values of 300 and 1500 is the increment in which the MATLAB script will process the population. Essentially, the script will first evaluate the cost associated with 300 substitutes, then 400, continuing in increments of 100 until the pool size reaches 1500. In order to determine an accurate pool size of substitutes, as well as the associated cost, the simulation must be run over varying intervals and iterations. The parameters of each simulation, as well as the resulting graph, are presented below.

    I. Optimal Pool Size for Substitutes assuming all Non-Degreed Substitute Teachers in Pool and All Full Time Teachers with One Year Experience utilized to cover shortfalls.
   a. Daily Pay Rate of Substitutes (p = 70)
        b. Daily Pay Rate of Full Time Teachers with Per Diem (r = 299.41)
        c. 1,000 Simulations Run on the Interval 300:100:1500
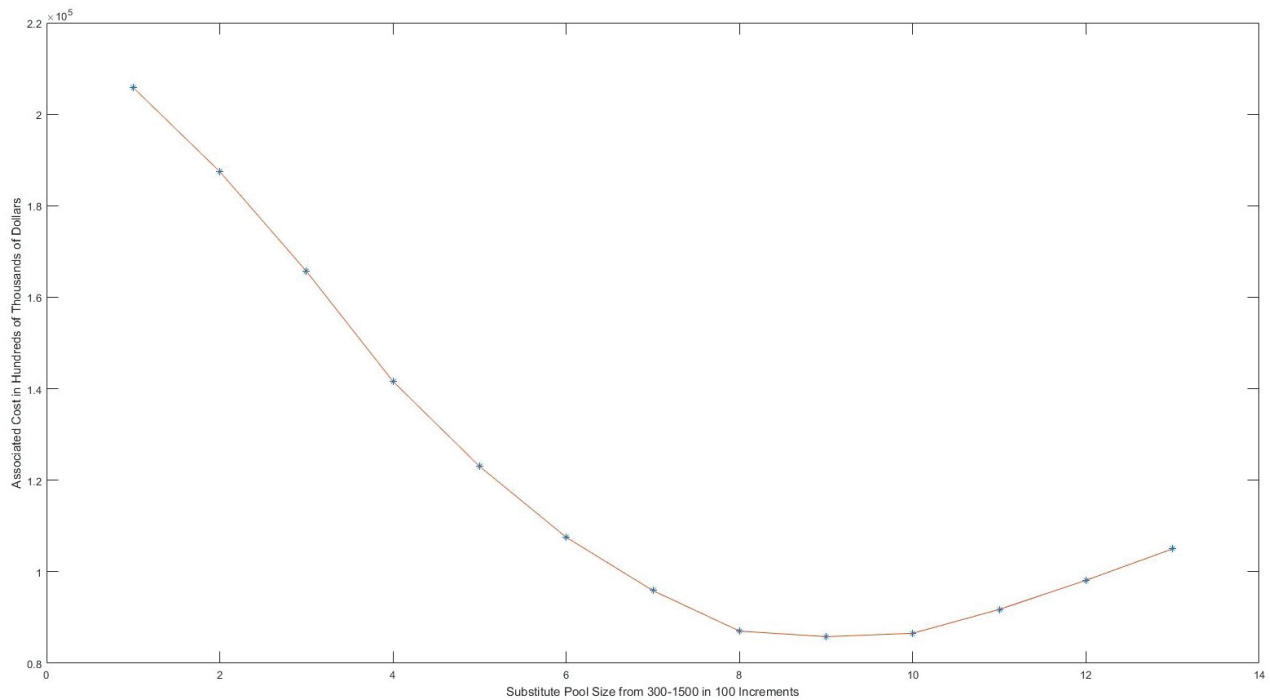
From the initial simulation it is determined that the optimal pool size is between 900 and 1200 substitute teachers. The next simulation narrows the range of the optimal pool size considerably.

    II. Second Parameterization for Initial Simulation of Degreed Substitutes and One Year Teachers.
        a. Daily Pay Rates Unchanged
   b. 100,000 Simulations Run on the Interval 900:25:1200
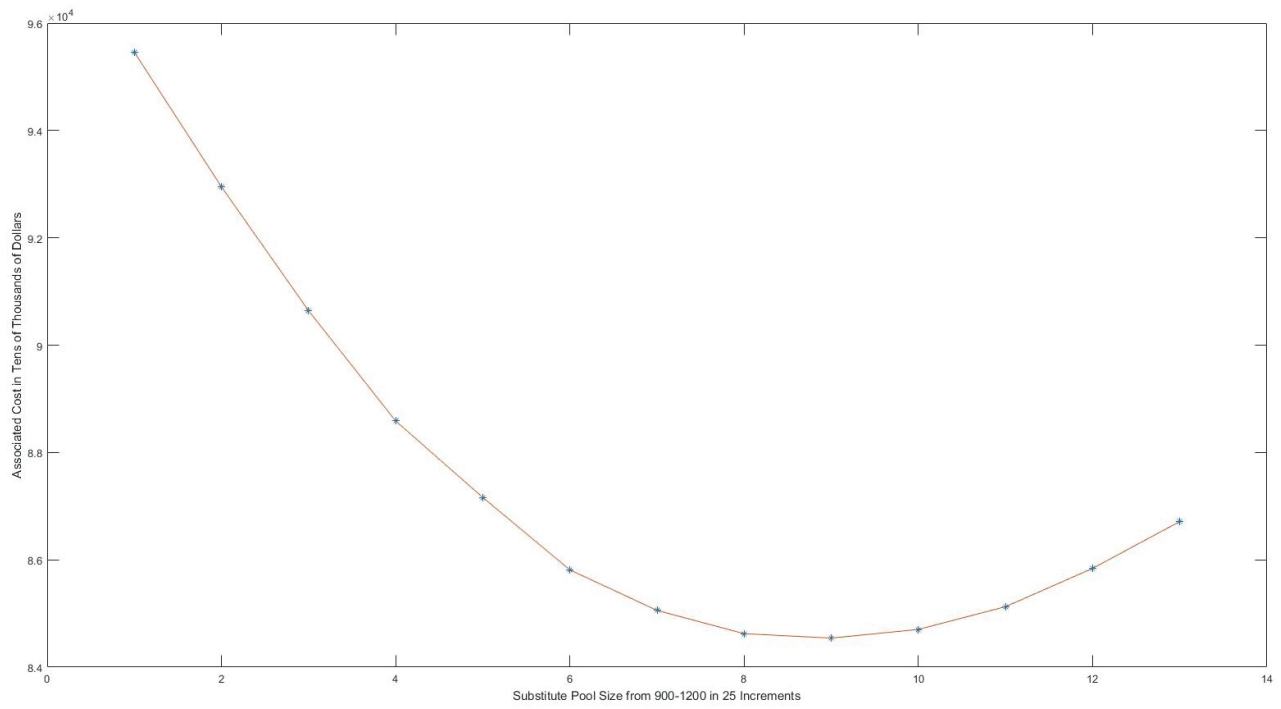
The second simulation narrowed the range of the optimal pool size of substitutes to between 1050 and 1125. A third simulation narrows the range of the optimal pool size even further.

    III. Third Parameterization for Degreed Substitutes and One Year Teachers

        a. Daily Pay Rates Unchanged

   b. 19,500,000 Simulations Run on the Interval 1050:5:1125
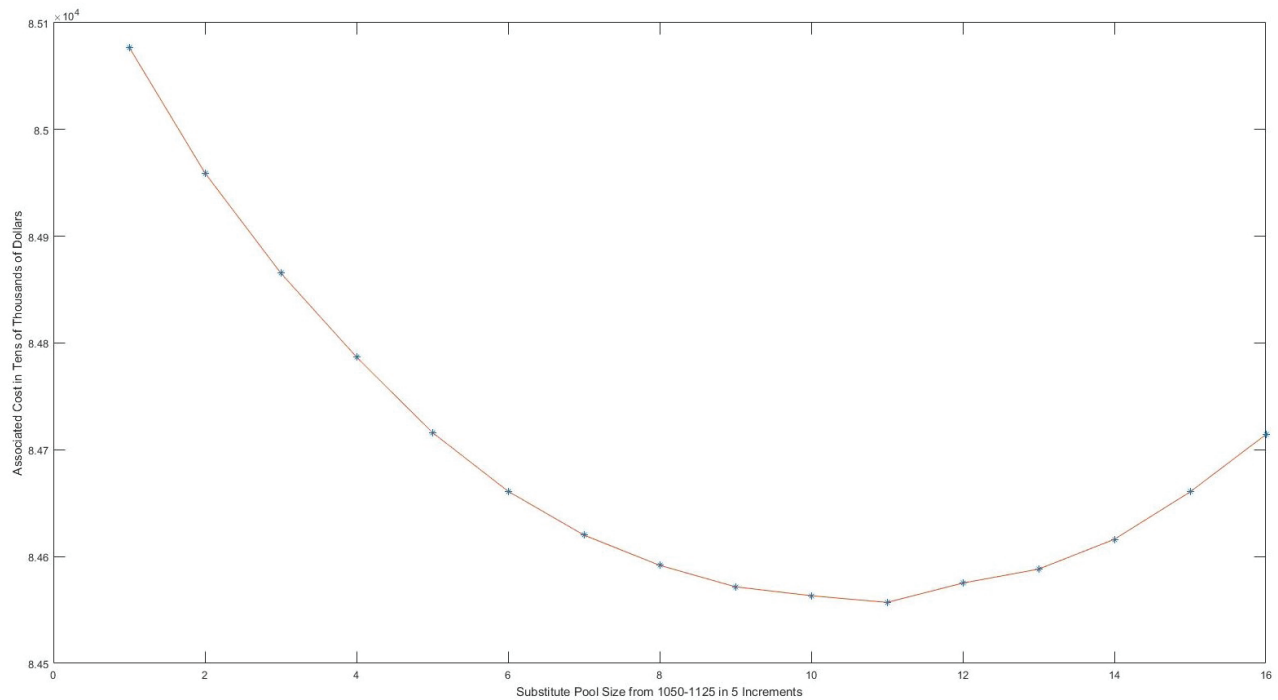
The third simulation narrowed the size of the pool to between 1095 and 1105 substitutes. This fourth and final simulation—based upon the author's script and the raw data provided by the district—determines the optimal substitute teacher range with the greatest accuracy.

IV. Fourth Parameterization for Degreed Substitutes and One Year Teachers
    a. Daily Pay Rates Unchanged
    b. 10,000,000,000 Simulations Run on the Interval 1095:1:1105
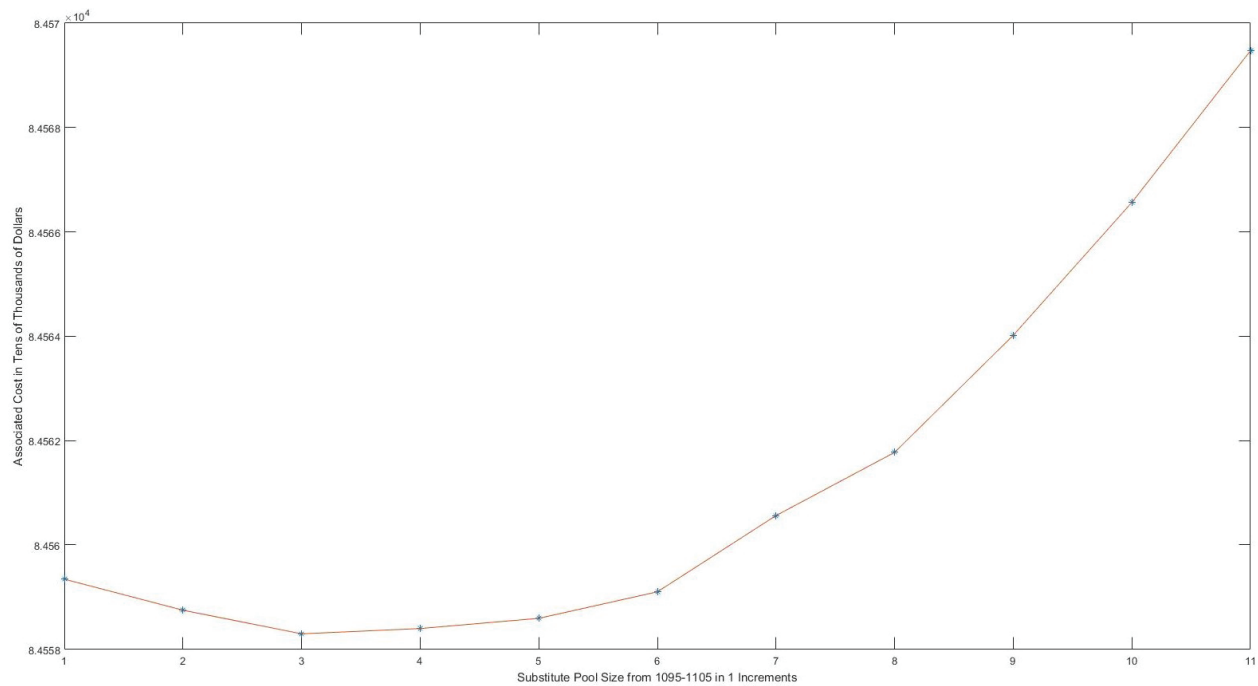
This fourth and final simulation has resulted in a highly accurate pool size that minimizes cost. The optimal pool size is 1097 Substitute Teachers, which results in a total cost of approximately 84,559 dollars. However, it must be reiterated that this is based upon the assumption that all Substitutes do not yet have a College Degree, and all Teachers assigned to fill any shortages are Teachers with One Year of Experience. Furthermore, it is assumed that all One Year Teachers will only be required to cover one additional class per day in addition to their full-time schedule.

The model can be adjusted to account for varying pay rates of substitutes as well as teachers with different levels of experience. Adjusting these parameters will change the optimal pool size for substitutes as well as their associated cost. The process that is utilized in order to achieve these models is similar to the process used to determine the initial model. The individual preparing the model need only adjust the pay rates of the Substitute Teachers and Full Time Teachers and the interval upon which the pool size is evaluated. To demonstrate the efficacy of variations in this modeling process, the final results of the remaining combinations of personnel and the associated cost are presented below. Also presented are the final results of all the initial graphical representations achieved with their corresponding parameters

Qualifications of Instructors in Pool	Daily Pay Rate of Substitutes	Daily Pay Rate of Full Time Teachers with Per Diem	Substitute Pool Population Interval	Number of Simulations	Optimal Pool Size of Substitutes	Total Cost to Denton Independent School District
Non-Degreed Substitutes, One Year Teachers	70	299.41	1095:1:1105	10,000,000,000	1097	84,559
Non-Degreed Substitutes, Ten Year Teachers	70	314.25	1105:1:1115	10,000,000,000	1107	84,927
Degreed Substitutes, One Year Teachers	80	299.41	1070:1:1080	10,000,000,000	1073	95,414
Degreed Substitutes, Ten Year Teachers	80	314.25	1075:1:1085	10,000,000,000	1083	95,868
Certified Substitutes, One Year Teachers	90	299.41	1050:1:1060	10,000,000,000	1057	106,062
Certified Substitutes, Ten Year Teachers	90	314.25	1055:1:1065	10,000,000,000	1063	106,590

Conclusions

The models derived from the data provided by the school district demonstrate that the optimal pool size for substitute teachers varies based upon respective pay rates. Should the pay rate for the substitutes increase, while the pay rate for full-time teachers remains unchanged, the optimal pool size shall decrease. As the pay rate for each full-time teacher increases, while the pay rate of substitutes remains unchanged, the optimal pool size shall increase. While these parameters do result in a shift of the optimal pool size, this shift is not extreme in either a positive or negative direction.

Through an interpretation of the obtained models, an overall pool size of substitutes between 1057 and 1107 should meet the needs of the school district under all expected circumstances. This pool size provides the needed coverage for absent teachers while simultaneously minimizing the overall cost to DISD. However, this is a general conclusion to cover most conceivable situations. In order to determine an optimal pool size with specific parameters, the necessary model must be defined and evaluated.

References

Denton Independent School District Human Resources. (2017). *Payroll Adjustment Schedule.* Retrieved April 11, 2018, from Denton ISD: http://www.dentonisd.org/cms/lib/TX21000245/Centricity/Domain/19/17-18%20Payroll%20Adjustment%20Schedule%202.1.18.pdf

Fox, W.P., Giordano, F.R., & Horton, S.B. (2014). *A First Course in Mathematical Modeling.* Australia: Brooks/Cole, Cengage Learning.