# EDUC 545 Final Paper

Juan Manubens

University Of Pennsylvania, Philadelphia PA 19104, USA,
manubens@wharton.upenn.edu,
https://github.com/juanmanubens

**Abstract.** This paper aims to (i) replicate an existing working paper by Reardon et al [1]. using machine learning algorithms that are relatively common in the realm of predictive modeling but have not been actively used in the field of education research and (ii) show how this procedure can be replicated with relative ease with R, a functional programming language used for statistical computing and graphics.

**Keywords:** Machine Learning, Regularization, Lasso, Ridge, Automatic Feature Selection, Subset Selection, Algorithms, Computational Social Science, Achievement Gaps, Inequality, Descriptive Analysis

## 1 Introduction

### 1.1 Motivation

The original paper estimated racial/ethnic achievement gaps in several hundred metropolitan areas and several thousand school districts in the United States using novel data from 200 million standardized math and reading tests administered to elementary and middle school students from 2009-2012. The paper found considerable substantial geographic variation in the magnitude of achievement gaps, ranging from nearly 0 in some places to larger than 1.2 standard deviations in others. A vector of economic, demographic, segregation and schooling characteristics variables explained roughly three-quarters of the geographic variation in these gaps, and the paper found the strongest correlates of achievement gaps to be racial/ethnic differences in parental income, parental education, and racial/ethnic segregation.

This paper will attempt to replicate these results using 2 different statistical approaches.

### 1.2 Data and Analysis Plan

Our data contains two dependent variables $Y_1 = gapblk\_adj; Y_2 = gaphsp\_adj$. For each $Y$, 80 predictors, $X_p$, are considered. These are split into 4 categories:

- **A: Socioeconomic and Family Structure** *- 23 variables considered*
- **B: Segregation** *- 10 variables considered*
- **C: Racial/Ethnic Composition** *- 17 variables considered*
- **D: School Quality** *- 30 variables considered*

The procedure used in this paper is as follows:

**1. Data Preprocessing**

1. The data frames containing the raw data are merged
2. The joined data frame is split into two new dataframes:
   (a) ***gapblk:*** data frame with non-missing values of *gapblk_adj*
   (b) ***gaphsp:*** data frame with non-missing values of *gaphsp_adj*

**2. Data Weighting**

1. Generate variable weights for each data set: $w_i$ : weight $= \frac{1}{SE(Y_i)^2}$
   (a) Both weights are kept as separate vectors in the code, as $W_{gapblk}$ and $W_{gaphsp}$
2. Apply variable weights to the data: $X_{wtd,i} = X_i \cdot \sqrt{w_i}$
   (a) The separate datasets are copied into new data frames, to which the weights are applied
3. Demean the data to account for state fixed effects: $X_{cen,wtd,i} = X_{wtd,i} - \bar{X}_{wtd,j}$
   (a) A double *for* loop is used to apply demeaning to all $X_i$'s for each state $j$

## 3. Modeling and Analysis

1. ***Ordinary Least Squares Fitting - Exhaustive Selection Models***
   (a) An exhaustive subset selection algorithm selects a subset of the variables to be used for regression for each predictor category.
   (b) The output of all categories is combined into a final full set, with which the algorithm is reapplied to select the final full model or OLS regression for each $Y$.
   (c) The algorithm uses 3 alternative criteria:
      i. Minimum *Mallow's* $C_p = \frac{1}{n}(RSS + 2d\hat{\alpha}^2)$
      ii. Minimum *Bayesian Information Criterion*, $BIC = \frac{1}{n}(RSS + log(n)d\hat{\alpha}^2)$
      iii. Maximum *Adjusted* $R^2 = 1 - \frac{RSS/(n-d-1)}{TSS/(n-1)}$
2. ***Regularized Fitting - Elastic Net Models***
   (a) We fit a model involving all $X_p$ predictors, but the estimated coefficients are shrunken towards zero relative to the OLS estimates.
   (b) This regularization reduces variance and can also perform variable selection.
3. ***Results***
   (a) Model Coefficients and Statistics such as $R^2$ and $RMSE$ are extracted.
   (b) Results are analyzed and performance is directly compared between the full OLS models and the best models extracted from the Elastic Net models.

The R code is excluded from this paper, but the full results can be found in Section 2 and in the Appendix.

## 2 Statistical Modeling

For each data frame containing $Y_1 = gapblk\_adj; Y_2 = gaphsp\_adj$, we test both statistical methods. For OLS Regression, we apply an exhaustive subset selection algorithm to select the optimal variables for regression within each category, and then reapply the method to the combined selection of all categories. For Elastic Net, we shrink coefficients based on our choice of $\alpha$ and select an optimal *tuning parameter* from cross-validation (see Section 2.2).

### 2.1 Ordinary Least Squares Fitting - Exhaustive Selection Models

The optimal subset selection algorithm's procedure, as described by Tibshirani et al. [2], is as follows:

1. ***Best Subset Selection Algorithm***
   (a) Let $\mathcal{M}_0$ denote the *null model*, which contains no predictors. This model simply predicts the sample mean, $Y_i = \hat{\beta}_0$, for each $x_i$.
   (b) For $k = 1, 2, ...p$ :
      i. Fit all $\binom{p}{k}$ that contain exactly $k$ predictor's
      ii. Pick the best among the $\binom{p}{k}$ models, and denote this model as $\mathcal{M}_k$, which is the optimal model with minimum $RSS$ and maximum $R^2$
   (c) Select a single best model among $\mathcal{M}_0, ..., \mathcal{M}_p$ using cross-validated prediction error, $C_p$ (AIC), BIC, or adjusted $R^2$. These criteria minimize test error *(i,ii)* and maximize variance explained *(iii)* :
      i. Minimum *Mallow's* $C_p = \frac{1}{n}(RSS + 2d\hat{\alpha}^2)$
      ii. Minimum *Bayesian Information Criterion*, $BIC = \frac{1}{n}(RSS + log(n)d\hat{\alpha}^2)$
      iii. Maximum *Adjusted* $R^2 = 1 - \frac{RSS/(n-d-1)}{TSS/(n-1)}$

Once we have selected our variables we fit a model such that we can minimize RSS:

$$\min_{\beta_0,\beta} RSS = \sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2 \tag{1}$$

While in theory it would be optimal to simply use this algorithm with all 80 predictors, it is computationally impractical with a large $p$, considering that we would have to select among $2^p$ models. At $p = 80$, this results in $2^{80} \approx 1.209 \cdot 10^{24}$ models, and thus our current approach is the most practical given the computational constraints.

## OLS Fitting Results

| OLS Exhaustive Selection Models | BLK | | HSP | |
|---|---|---|---|---|
| Model | Adjusted RSQ | RMSE | Adjusted RSQ | RMSE |
| A: Socioeconomic and Family Structure | 0.9149 | 5.784 | 0.8039 | 4.137 |
| B: Segregation | 0.9154 | 5.767 | 0.8014 | 4.025 |
| C: Racial/Ethnic Composition | 0.8134 | 8.564 | 0.6936 | 5.175 |
| D: School Quality | 0.8291 | 8.196 | 0.6969 | 5.143 |
| All Variables | 0.9328 | 5.137 | 0.8525 | 3.588 |

| | All Variables | | | | | |
|---|---|---|---|---|---|---|
| | Category | Variable | Estimate | Std. Error | t value | Pr(>\|t\|) |
| blk | - | (Intercept) | 0.00 | 0.10 | 0.00 | 1.00 |
| | A | baplus_all | 0.54 | 0.03 | 17.72 | 0.00 |
| | | poverty517_all | -0.59 | 0.05 | -11.40 | 0.00 |
| | | giniall | 1.24 | 0.05 | 24.77 | 0.00 |
| | | incVblkwht | 0.11 | 0.01 | 13.02 | 0.00 |
| | B | diffexplch_blkwht | 0.85 | 0.05 | 17.44 | 0.00 |
| | | paredVblkwht | 0.18 | 0.01 | 17.80 | 0.00 |
| | C | pctmexico | 0.06 | 0.01 | 5.91 | 0.00 |
| | D | hsflnfl | -0.50 | 0.06 | -8.37 | 0.00 |

| | All Variables | | | | | |
|---|---|---|---|---|---|---|
| | Category | Variable | Estimate | Std. Error | t value | Pr(>\|t\|) |
| hsp | - | (Intercept) | 0.00 | 0.10 | 0.00 | 1.00 |
| | A | baplus_all | 0.59 | 0.04 | 14.27 | 0.00 |
| | | incVhspwht | 0.07 | 0.01 | 6.82 | 0.00 |
| | | occdiff_whthsp | 0.13 | 0.03 | 3.93 | 0.00 |
| | B | perell | 0.47 | 0.05 | 9.00 | 0.00 |
| | | flunch_all | -0.16 | 0.03 | -4.76 | 0.00 |
| | | diffexplch_hspwht | 0.59 | 0.09 | 6.91 | 0.00 |
| | | paredVhspwht | 0.23 | 0.01 | 16.78 | 0.00 |
| | D | charter_na | 0.23 | 0.03 | 7.15 | 0.00 |

Overall, there is a common hierarchy between both achievement gaps, achieving a very similar result to the original paper. Variables from Category A: Socioeconomic & Family Structure and Category B: Segregation seem to be the most predictive and least error-prone for both achievement gaps. There are common variables in both of the full models, such as *baplus_all*, % of adults with a Bachelor's degree or higher accross all races. Interestingly, no variables from Category C: Racial/Ethnic Composition, were selected for the Hispanic Gap Full OLS model.

| | Ordinary Least Squares Exhaustive Selection Models | |
|---|---|---|
| % Importance by Model | Black Gap | Hispanic Gap |
| A: Socioeconomic and Family Structure | 61.0% | 32.1% |
| B: Segregation | 25.2% | 58.5% |
| C: Racial/Ethnic Composition | 1.5% | 0.0% |
| D: School Quality | 12.3% | 9.4% |

For this study, variable importance was calculated as the proportion between the absolute value of a variable's coefficient and the sum of the absolute value of all of the model coefficients, or $\frac{|\beta_{j,Y}|}{\sum_{j=1}^{p} \beta_{j,Y}}$. For the black achievement gap, Socioeconomic variables, followed by Segregation variables, have the highest impact in the model. The same categories apply for the hispanic achievement gap, but in the reverse order.

### 2.2 Regularized Fitting - Elastic Net Models

As an alternative to OLS, we can fit a model containing all $p$ predictors using a technique that constrains or regularizes the coefficient estimates, or equivalently, that shrinks the $\beta_j$ estimates towards zero. This regularization reduces variance and can perform variable selection.There are 3 ways to approach regularization: *(i) Ridge Regression (ii) Lasso (iii) Hybrid Models between Ridge and Lasso, also known as Elastic Net.[2][3]*

**(i) Ridge Regression** We fit *Ridge* coefficients, $\hat{\beta_\lambda^R}$, such that we can minimize RSS and an additional *shrinkage penalty*:

$$\min_{\beta_0,\beta} RSS = \sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \tag{2}$$

*Ridge* regression penalizes the magnitude of the coefficients, not not the complexity of the model, and this is it's main disadvantage: it will include all $p$ predictors in the final model. This is why *Lasso* is a good alternative.

**(ii) Lasso** We fit *Lasso* coefficients, $\hat{\beta_\lambda^L}$, such that we can minimize RSS and an additional *shrinkage penalty*:

$$\min_{\beta_0,\beta} RSS = \sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \mid \beta_j \mid \tag{3}$$

Unlike *Ridge*, *Lasso* performs variable selection, yielding only a subset of the variables.

**(iii) Elastic Net** We make a hybrid between *Ridge* and *Lasso* through the use of the parameter $\alpha$, where $0 \leq \alpha \leq 1$. When $\alpha = 0$, we fit a Ridge Model, and when $\alpha = 1$, we fit a Lasso Model:

$$\min_{\beta_0,\beta} RSS = \sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \left[ (1-\alpha)\lambda \sum_{j=1}^{p} \beta_j^2 + (\alpha)\lambda \sum_{j=1}^{p} \mid \beta_j \mid \right] \tag{4}$$

*Lambda*, $\lambda \geq 0$, is a *tuning parameter* that is determined through cross-validation. In this study, our criteria is the *"one-standard-error"* rule when selecting the best model, as suggested by Tibshirani et al. This acknowledges estimation error and is more parsimonious: it helps us choose the simplest model whose accuracy is comparable with the best model [3][4].

**Regularized Fitting Results**

| Elastic Net Models | BLK | | HSP | |
|---|---|---|---|---|
| Model | RSQ | RMSE | RSQ | RMSE |
| Ridge | 0.9329 | 5.135 | 0.8612 | 3.478 |
| Alpha = 0.2 | 0.9357 | 5.026 | 0.8509 | 3.605 |
| Alpha = 0.4 | 0.9346 | 5.069 | 0.8529 | 3.581 |
| Alpha = 0.6 | 0.9338 | 5.101 | 0.8522 | 3.590 |
| Alpha = 0.8 | 0.9345 | 5.071 | 0.8528 | 3.583 |
| Lasso | 0.9352 | 5.045 | 0.8500 | 3.617 |

The best performing models for both achievement gaps were closer to *Ridge* than to *Lasso*. For the black achievement gap, the hybrid *Elastic Net* model with $\alpha = 0.2$ performed best, and for the hispanic achievement gap, the *Ridge* model performed best.

| | Elastic Net - Regularization | |
|---|---|---|
| | Alpha at 0.2 | Ridge |
| % Importance by Model | Black Gap | Hispanic Gap |
| A: Socioeconomic and Family Structure | 39.9% | 31.6% |
| B: Segregation | 29.2% | 23.9% |
| C: Racial/Ethnic Composition | 11.9% | 10.2% |
| D: School Quality | 18.9% | 34.3% |

For this study, variable importance was calculated as the proportion between the absolute value of a variable's coefficient and the sum of the absolute value of all of the model coefficients, or $\frac{|\beta_{j,Y}|}{\sum_{j=1}^{p} \beta_{j,Y}}$. Similar to the OLS models, Socioeconomic variables have a very high impact in the model. Nevertheless, the model gives more weight to Demographic and Schooling variables. For the black achievement gap, Socioeconomic variables, followed by Segregation variables, have the highest impact in the model. For the hispanic achievement gap, Schooling variables have the highest impact based on this study's metric.

**Model Comparison & Conclusion**

| % Importance by Model | Ordinary Least Squares |  | Elastic Net - Regularization |  |
|---|---|---|---|---|
|  | Exhaustive Selection Models |  | Alpha at 0.2 | Ridge |
|  | Black Gap | Hispanic Gap | Black Gap | Hispanic Gap |
| A: Socioeconomic and Family Structure | 61.0% | 32.1% | 39.9% | 31.6% |
| B: Segregation | 25.2% | 58.5% | 29.2% | 23.9% |
| C: Racial/Ethnic Composition | 1.5% | 0.0% | 11.9% | 10.2% |
| D: School Quality | 12.3% | 9.4% | 18.9% | 34.3% |

For the black achievement gap, the *Elastic Net* model with $\alpha = 0.2$ reached the same conclusion as the OLS Full Model in terms of variable importance and respective rank, with Socioeconomic variables, followed by Segregation variables being the most important and having the highest impact in the model, accordingly. However, the Ridge Model and the OLS Model for the hispanic gap do not reach a consensus, as seen above. Schooling variables are given considerable weight in the *Ridge* model.

Overall, there is merit to using alternative fitting procedures such as *Ridge*, *Lasso*, and *Elastic Net*, in conjunction with traditional methods such as Ordinary Least Squares. It can provide a reasonable reality check through comparison and contrast, and differences between models can raise new research questions as well as provide a different narrative than what we would get from Least Squares. It is also easily reproducible, as I will show in the code, separate from this document. Finally, the use of Exhaustive Subset Selection Algorithm is also useful tool to allow researchers to select a relatively optimal subset of variables with minimal selection bias, helping reach good results at a very practical speed conditional to computational constraints.

## References

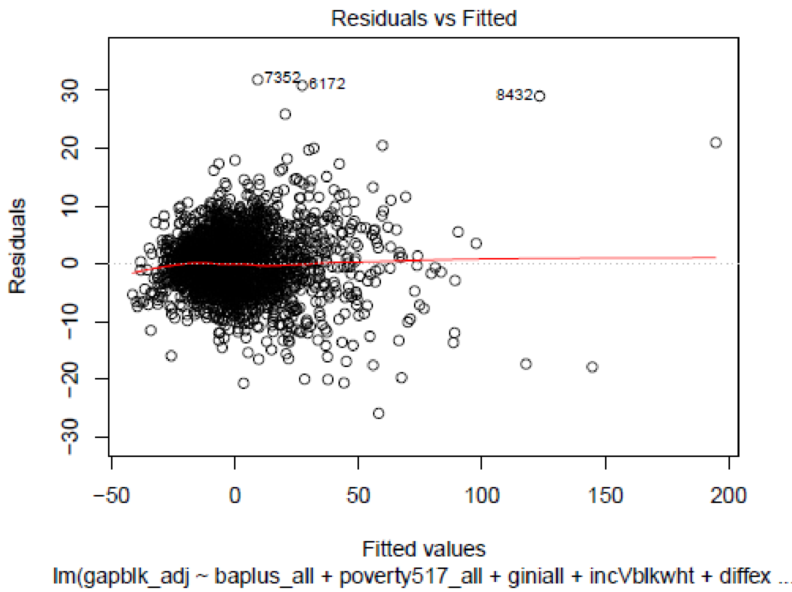

1. Reardon, S.F., Kalogrides, D., & Shores, K. (2016). The Geography of Racial/Ethnic Test Score Gaps (CEPA Working Paper No.16-10). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp16-10
2. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. (2014). An Introduction to Statistical Learning: With Applications in R. Springer Publishing Company, Incorporated.
3. Hastie T, Tibshirani R, Friedman J. (2009). The Elements of Statistical Learning: data mining, inference, and prediction, 2nd Edition. Springer Publishing Company, Incorporated., New York

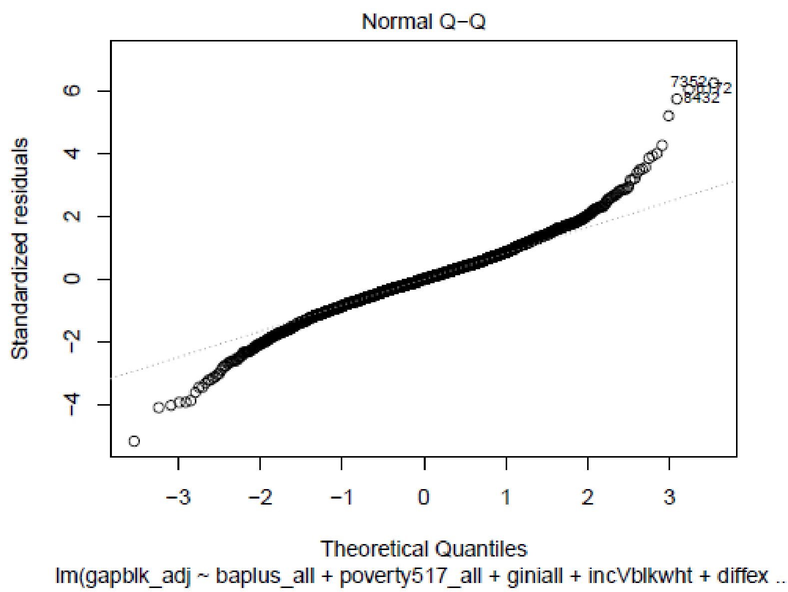
4. Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. Journal of Cheminformatics, 6(1), 10. http://doi.org/10.1186/1758-2946-6-10

# 3 Appendix

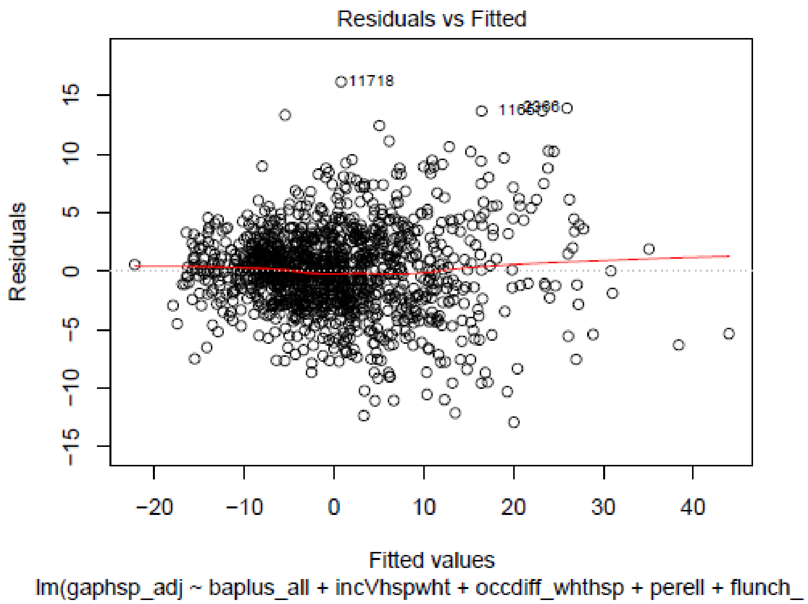
## 3.1 Ordinary Least Squares Fitting - Exhaustive Selection Models

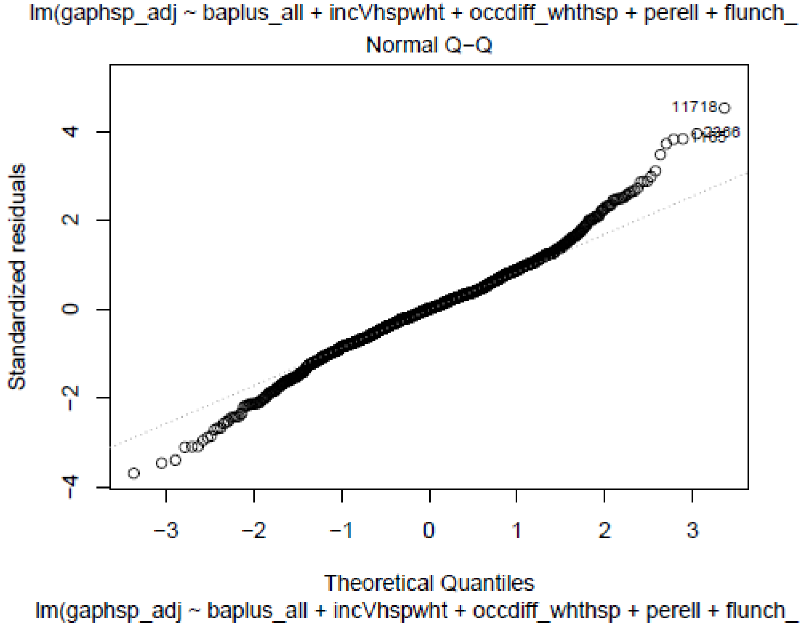
### 3.1.1 OLS Model Plots - Y: gapblk_adj



### 3.1.2 OLS Model Plots - Y: gaphsp_adj



## 3.2 Regularized Fitting - Elastic Net Models

### 3.2.1 Elastic Net Model Coefficients for Category A: Socioeconomic & Family Structure

| | Category A: Socioeconomic and Family Structure | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | Intercept | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | profocc_all | 0.24 | 0.24 | 0.21 | 0.18 | 0.17 | 0.15 | 1 | 1 | 1 | 1 | 1 | 1 |
| | baplus_all | 0.20 | 0.38 | 0.43 | 0.47 | 0.51 | 0.52 | 1 | 1 | 1 | 1 | 1 | 1 |
| | poverty517_all | -0.09 | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | singmom_all | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | snap_all | -0.10 | -0.10 | -0.05 | -0.02 | -0.03 | -0.05 | 1 | 1 | 1 | 1 | 1 | 1 |
| | rent_all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | samehouse_all | 0.03 | 0.08 | 0.09 | 0.10 | 0.13 | 0.13 | 1 | 1 | 1 | 1 | 1 | 1 |
| | unemp_all | -0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | inc50all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 0 |
| | incrat9010all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| blk | incrat9050all | 0.01 | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 | 1 | 1 | 1 | 1 | 1 | 1 |
| | incrat5010all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | giniall | 0.07 | 0.12 | 0.10 | 0.05 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | incVblkwht | 0.05 | 0.05 | 0.06 | 0.06 | 0.06 | 0.06 | 1 | 1 | 1 | 1 | 1 | 1 |
| | hhsize_all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | povdiff517_whtblk | 0.05 | 0.01 | 0.01 | 0.00 | 0.01 | 0.01 | 1 | 1 | 1 | 1 | 1 | 1 |
| | unempdiff_whtblk | 0.17 | 0.07 | 0.03 | 0.01 | 0.02 | 0.03 | 1 | 1 | 1 | 1 | 1 | 1 |
| | occdiff_whtblk | 0.16 | 0.11 | 0.11 | 0.11 | 0.11 | 0.10 | 1 | 1 | 1 | 1 | 1 | 1 |
| | singmomdiffwhtblk | 0.07 | 0.06 | 0.05 | 0.05 | 0.05 | 0.05 | 1 | 1 | 1 | 1 | 1 | 1 |
| | rentdiffwhtblk | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.07 | 1 | 1 | 1 | 1 | 1 | 1 |
| | samehousediffwhtblk | 0.05 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 1 | 1 | 1 | 1 | 1 | 1 |
| | medvalue | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | medrent | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |

| | Category A: Socioeconomic and Family Structure | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | (Intercept) | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | profocc_all | 0.24 | 0.26 | 0.24 | 0.21 | 0.19 | 0.14 | 1 | 1 | 1 | 1 | 1 | 1 |
| | baplus_all | 0.24 | 0.29 | 0.35 | 0.38 | 0.42 | 0.45 | 1 | 1 | 1 | 1 | 1 | 1 |
| | poverty517_all | -0.11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | singmom_all | -0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | snap_all | -0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | rent_all | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | samehouse_all | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | unemp_all | -0.33 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | inc50all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | incrat9010all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| hsp | incrat9050all | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 0 | 0 |
| | incrat5010all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | giniall | 0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | incVhspwht | 0.05 | 0.06 | 0.05 | 0.05 | 0.05 | 0.05 | 1 | 1 | 1 | 1 | 1 | 1 |
| | hhsize_all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | povdiff517_whthsp | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | unempdiff_whthsp | 0.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | occdiff_whthsp | 0.15 | 0.15 | 0.13 | 0.11 | 0.10 | 0.09 | 1 | 1 | 1 | 1 | 1 | 1 |
| | singmomdiffwhthsp | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | rentdiffwhthsp | 0.10 | 0.10 | 0.09 | 0.09 | 0.09 | 0.08 | 1 | 1 | 1 | 1 | 1 | 1 |
| | samehousediffwhthsp | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | medvalue | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | medrent | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |

### 3.2.2 Elastic Net Model Coefficients for Category B: Segregation

| | Category B: Segregation | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | perfrl | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perell | 0.13 | 0.07 | 0.04 | 0.02 | 0.02 | 0.03 | 1 | 1 | 1 | 1 | 1 | 1 |
| | perspeced | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | frl | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| blk | ratstutch_whtblk | 0.03 | 0.06 | 0.08 | 0.10 | 0.10 | 0.11 | 1 | 1 | 1 | 1 | 1 | 1 |
| | flunch_all | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | diffexplch_blkwht | 0.39 | 0.59 | 0.57 | 0.55 | 0.59 | 0.61 | 1 | 1 | 1 | 1 | 1 | 1 |
| | paredVblkwht | 0.15 | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 1 | 1 | 1 | 1 | 1 | 1 |
| | snapdiffwhtblk | 0.12 | 0.11 | 0.11 | 0.10 | 0.10 | 0.10 | 1 | 1 | 1 | 1 | 1 | 1 |
| | privdiff_whtblk | 0.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |

| | Category B: Segregation | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | perfrl | -0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perell | 0.30 | 0.26 | 0.26 | 0.25 | 0.25 | 0.22 | 1 | 1 | 1 | 1 | 1 | 1 |
| | perspeced | 0.15 | 0.08 | 0.07 | 0.05 | 0.05 | 0.01 | 1 | 1 | 1 | 1 | 1 | 1 |
| | frl | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| hsp | ratstutch_whthsp | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 0 | 0 |
| | flunch_all | -0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | diffexplch_hspwht | 0.42 | 0.45 | 0.43 | 0.40 | 0.39 | 0.35 | 1 | 1 | 1 | 1 | 1 | 1 |
| | paredVhspwht | 0.12 | 0.16 | 0.18 | 0.20 | 0.21 | 0.22 | 1 | 1 | 1 | 1 | 1 | 1 |
| | snapdiffwhthsp | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | privdiff_whthsp | 0.15 | 0.07 | 0.06 | 0.04 | 0.04 | 0.01 | 1 | 1 | 1 | 1 | 1 | 1 |

### 3.2.3 Elastic Net Model Coefficients for Category C: Racial/Ethnic Composition

| | Category C: Racial/Ethnic Composition | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | perind | -0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perasn | -0.11 | -0.18 | -0.15 | -0.14 | -0.16 | -0.18 | 1 | 1 | 1 | 1 | 1 | 1 |
| | perblk | 0.09 | 0.08 | 0.06 | 0.05 | 0.05 | 0.06 | 1 | 1 | 1 | 1 | 1 | 1 |
| | perwht | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | ind | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | asn | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | blk | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| blk | wht | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 0 | 1 |
| | pctenglish1 | 0.02 | 0.04 | 0.04 | 0.04 | 0.05 | 0.04 | 1 | 1 | 1 | 1 | 1 | 1 |
| | pctenglish2 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctenglish3 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctforeign | 0.08 | 0.06 | 0.06 | 0.06 | 0.06 | 0.05 | 1 | 1 | 1 | 1 | 1 | 1 |
| | pctmexico | 0.04 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 1 | 1 | 1 | 1 | 1 | 1 |
| | pctpuerto | 0.02 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | pctcuba | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctcentral | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | pctsouth | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |

| | Category C: Racial/Ethnic Composition | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | perind | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perasn | -0.07 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perhsp | 0.07 | 0.02 | 0.01 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 0 | 0 | 0 |
| | perwht | -0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | ind | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | asn | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | hsp | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| hsp | wht | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctenglish1 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctenglish2 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctenglish3 | -0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctforeign | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctmexico | 0.03 | 0.03 | 0.02 | 0.02 | 0.01 | 0.01 | 1 | 1 | 1 | 1 | 1 | 1 |
| | pctpuerto | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctcuba | -0.17 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctcentral | -0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pctsouth | -0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |

### 3.2.4 Elastic Net Model Coefficients for Category D: School Quality

| | Category D: School Quality | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | avgrdall | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | avgrdwht | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | avgrdblk | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | totenrl | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | nsch | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | ncharters | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | elmtch | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | tottch | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | aides | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | corsup | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | elmgui | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | stutch_all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | hswhtblk | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | hsflnfl | 0.01 | -0.08 | -0.04 | -0.01 | -0.05 | -0.10 | 1 | 1 | 1 | 1 | 1 | 1 |
| blk | ppexp_tot | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | ppexp_inst | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | pprev_tot | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | percharter_all | -0.14 | -0.14 | -0.13 | -0.13 | -0.13 | -0.13 | 1 | 1 | 1 | 1 | 1 | 1 |
| | totppe_fleslope | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | gslo | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | gshi | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | teenbirth_all | -0.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perfst_all | -0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perfstdiff_whtblk | 0.33 | 0.22 | 0.18 | 0.16 | 0.17 | 0.17 | 1 | 1 | 1 | 1 | 1 | 1 |
| | perabs_all | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perabsdiff_whtblk | 0.16 | 0.08 | 0.04 | 0.01 | 0.03 | 0.04 | 1 | 1 | 1 | 1 | 1 | 1 |
| | percharter | 0.16 | 0.02 | 0.00 | 0.00 | 0.04 | 0.05 | 1 | 1 | 1 | 1 | 1 | 1 |
| | charter_cr | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | charter_na | 0.04 | 0.05 | 0.04 | 0.04 | 0.03 | 0.03 | 1 | 1 | 1 | 1 | 1 | 1 |
| | chardiff_blk | -0.08 | -0.06 | -0.05 | -0.05 | -0.05 | -0.06 | 1 | 1 | 1 | 1 | 1 | 1 |

| | Category D: School Quality | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso | Ridge | Alpha_0.2 | Alpha_0.4 | Alpha_0.6 | Alpha_0.8 | Lasso |
| | avgrdall | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | avgrdwht | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | avgrdhsp | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | totenrl | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | nsch | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | ncharters | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | elmtch | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | tottch | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | aides | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | corsup | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | elmgui | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | stutch_all | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | hswhthsp | 0.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | hsflnfl | -0.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| hsp | ppexp_tot | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 0 | 0 |
| | ppexp_inst | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | pprev_tot | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 1 | 1 | 1 |
| | percharter_all | 0.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | totppe_fleslope | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | gslo | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | gshi | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 1 | 0 | 0 | 0 |
| | teenbirth_all | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perfst_all | -0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perfstdiff_whthsp | 0.76 | 0.09 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | perabs_all | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | perabsdiff_whthsp | 0.47 | 0.07 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 1 | 0 | 0 | 0 | 0 |
| | percharter | -0.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |
| | charter_cr | 0.05 | 0.02 | 0.02 | 0.02 | 0.02 | 0.00 | 1 | 1 | 1 | 1 | 1 | 0 |
| | charter_na | 0.03 | 0.02 | 0.03 | 0.04 | 0.05 | 0.07 | 1 | 1 | 1 | 1 | 1 | 1 |
| | chardiff_hsp | -0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1 | 0 | 0 | 0 | 0 | 0 |

### 3.2.5 Elastic Net Optimal Model Plots - Fitted versus Actual Values