Fairness Analysis Figures and Tables

Ariane Stark and Dasuni Jayawardena

Data Editing

Important Values

```
#race
cor(x=all_the_data_pull_cutoff$`total cases`,
    y=all_the_data_pull_cutoff$Total_Black/
      all_the_data_pull_cutoff$Total_Pop,
    method="spearman")
```

```
## [1] 0.4329403
```

```
cor.test(x=all_the_data_pull_cutoff$`total cases`,
    y=all_the_data_pull_cutoff$Total_Black/
      all_the_data_pull_cutoff$Total_Pop,
    method="spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  all_the_data_pull_cutoff$`total cases` and all_the_data_pull_cutoff$Total_Black/all_the_data_p
## S = 2862120159, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.4329403
```

```
#poverty
x_case <- all_the_data_pull_cutoff$`total cases`
y_pov <- all_the_data_pull_cutoff$Total_Households_Below_Poverty/
      (all_the_data_pull_cutoff$Total_Households_Above_Poverty+
         all_the_data_pull_cutoff$Total_Households_Below_Poverty)

cor(x=x_case[-c(1814)],
    y=y_pov[-c(1814)],
    method="spearman")
```

```
## [1] -0.08459665
```

```r
cor.test(x=x_case[-c(1814)],
    y=y_pov[-c(1814)],
    method="spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  x_case[-c(1814)] and y_pov[-c(1814)]
## S = 5.469e+09, p-value = 2.258e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.08459665
```

MAE Wilcoxon Tests

Wilcoxon rank sum between proportion of black constituent quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_mae and quart_2_mae
## W = 178316, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_mae and quart_3_mae
## W = 102559, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_mae and quart_4_mae
## W = 126617, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_3_mae and quart_4_mae
## W = 326137, p-value = 0.01052
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 1 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
```

```
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_2
## W = 31550, p-value = 0.06056
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_3
## W = 14316, p-value = 8.157e-05
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_4
## W = 16146, p-value = 6.584e-05
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 2 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_2
## W = 21654, p-value = 0.03436
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_3
## W = 9685, p-value = 3.063e-08
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_4
## W = 17708, p-value = 0.003986
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 3 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_2
## W = 12882, p-value = 0.01097
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_3
## W = 11604, p-value = 0.002071
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_4
## W = 10542, p-value = 0.05698
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 4 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_2
## W = 2195.5, p-value = 0.6978
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_3
## W = 4235, p-value = 0.2756
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_4
## W = 3159, p-value = 0.01546
## alternative hypothesis: true location shift is not equal to 0
```

RMAE Wilcoxon Tests

Wilcoxon rank sum between proportion of black constituent quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_rmae and quart_2_rmae
## W = 317414, p-value = 0.1259
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
```

```
##
## data:  quart_1_rmae and quart_3_rmae
## W = 320257, p-value = 0.06424
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_rmae and quart_4_rmae
## W = 334306, p-value = 0.0006008
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_3_rmae and quart_4_rmae
## W = 318906, p-value = 0.08115
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 1 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_2
## W = 33498, p-value = 0.4139
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_3
## W = 19176, p-value = 0.9022
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_4
## W = 23119, p-value = 0.255
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 2 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_2
## W = 26154, p-value = 0.2203
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_3
## W = 15890, p-value = 0.2776
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_4
## W = 24075, p-value = 0.01676
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 3 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_2
## W = 16912, p-value = 0.1077
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_3
## W = 15469, p-value = 0.269
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_4
## W = 14528, p-value = 0.002012
## alternative hypothesis: true location shift is not equal to 0
```

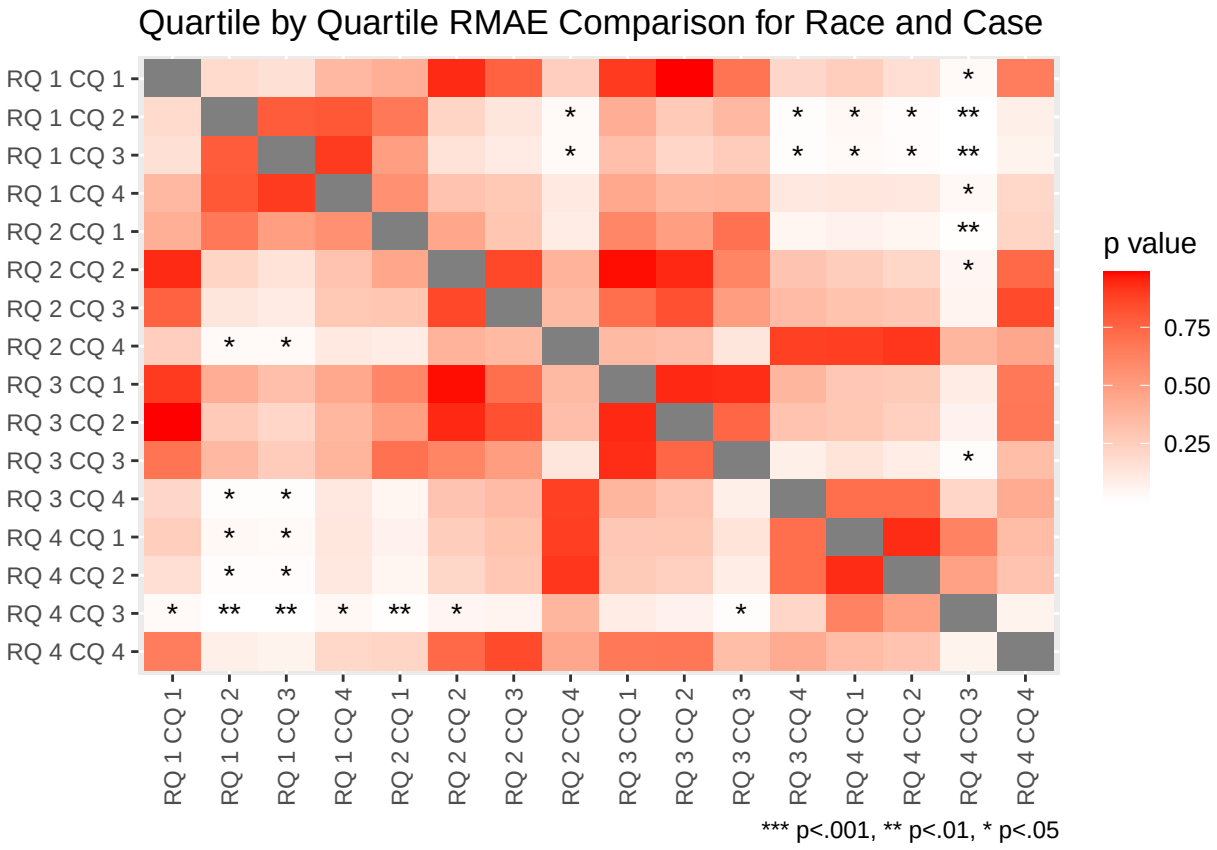
Within Covid Case Quartile 4 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_2
## W = 2484, p-value = 0.117
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_3
## W = 5635, p-value = 0.1231
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_4
## W = 4927, p-value = 0.2057
## alternative hypothesis: true location shift is not equal to 0
```

All 16 quartile quartile breakdowns for RMAE comparison



Summary Statistics Case Quartile

Summary Statistics By Race Quartile

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.118   5.912  10.279  17.038  17.890 574.088
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##    1.882   14.721   29.500   86.246   81.926 2559.765
```

Counts by Quartile Quartile Breakdown

```
##              Prop. Black Quartile
```

```
## Case Quartile lowest 25%  Q2  Q3 highest 25% Overall
##   lowest 25%         407 172  95         106     780
##   Q2                 214 229 139         198     780
##   Q3                 129 238 224         187     778
##   highest 25%         30 140 321         288     779
##   Overall            780 779 779         779    3117
```

Median MAE and RMAE tables

```
##              Prop. Black Quartile
## Case Quartile range                 lowest 25% Q2       Q3       highest 25%
##   lowest 25%  "[2,1.09e+03]"        "6.059"    "6.559"  "7.794"  "7.632"
##   Q2          "(1.09e+03,2.64e+03]" "14.103"   "14.882" "17.706" "15.147"
##   Q3          "(2.64e+03,6.82e+03]" "25.853"   "29.088" "29.603" "28.912"
##   highest 25% "(6.82e+03,1.25e+06]" "79.059"   "73.368" "91.324" "111.779"
##   Overall     ""                    "10.279"   "19.588" "35.618" "29.5"
##              Prop. Black Quartile
## Case Quartile Overall
##   lowest 25%  "6.647"
##   Q2          "14.912"
##   Q3          "28.515"
##   highest 25% "93.353"
##   Overall     "20.735"
```

MAE table generated by LaTex

		Prop. Black Quartile				
Case Quartile	Range	lowest 25%	Q2	Q3	highest 25%	Overall
lowest 25%	[2, 1.09e3]	6.059	6.559	7.794	7.632	6.647
Q2	(1.09e3, 2.64e3]	14.103	14.882	17.706	15.147	14.912
Q3	(2.64e3, 6.82e3]	25.853	29.088	29.603	28.912	28.515
highest 25%	(6.82e3, 1.25e6]	79.059	73.368	91.324	111.779	93.353
Overall		10.279	19.588	35.618	29.5	20.735

Table 1: Pink for statistically significant at $\alpha = 0.05$ and red for statistically significant at $\alpha = 0.00333$ (from the Bonferroni Correction over 15 tests) when comparing to Quartile 1 of Prop. Black. across the case quartiles and overall

RMAE Table code

```
##              Prop. Black Quartile
## Case Quartile range                 lowest 25% Q2      Q3      highest 25%
##   lowest 25%  "[2,1.09e+03]"        "0.898"    "0.901" "0.91"  "0.885"
##   Q2          "(1.09e+03,2.64e+03]" "0.906"    "0.887" "0.902" "0.887"
##   Q3          "(2.64e+03,6.82e+03]" "0.911"    "0.899" "0.905" "0.879"
##   highest 25% "(6.82e+03,1.25e+06]" "0.923"    "0.89"  "0.889" "0.894"
##   Overall     ""                    "0.903"    "0.896" "0.897" "0.887"
##              Prop. Black Quartile
## Case Quartile Overall
##   lowest 25%  "0.898"
```

```
##   Q2           "0.897"
##   Q3           "0.898"
##   highest 25% "0.891"
##   Overall      "20.735"
```

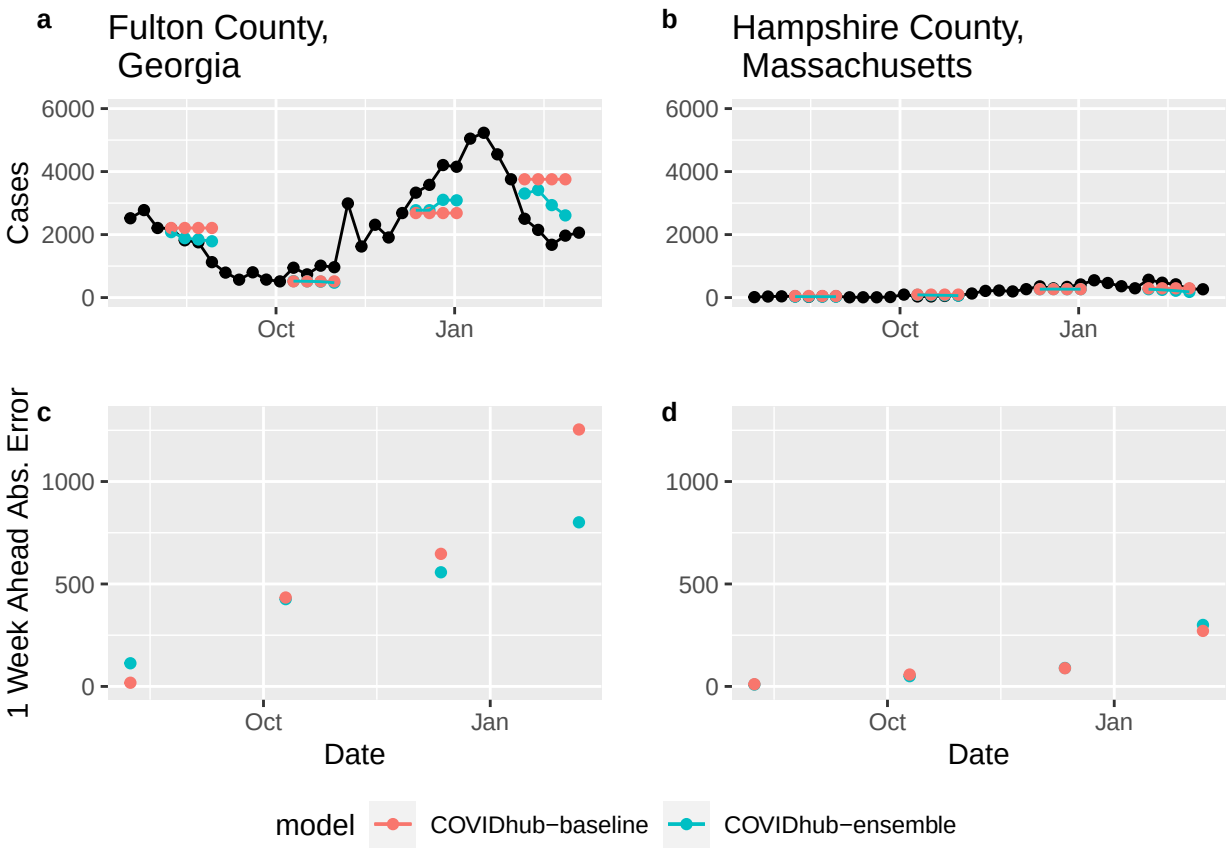
RMAE Table from LaTex

		Prop. Black Quartile				
Case Quartile	Range	lowest 25%	Q2	Q3	highest 25%	Overall
lowest 25%	[2, 1.09e3]	0.898	0.901	0.910	0.885	0.898
Q2	(1.09e3, 2.64e3]	0.906	0.887	0.902	0.887	0.897
Q3	(2.64e3, 6.82e3]	0.911	0.899	0.905	0.879	0.898
highest 25%	(6.82e3, 1.25e6]	0.923	0.890	0.889	0.894	0.891
Overall		0.903	0.896	0.897	0.887	0.896

Table 2: Pink for statistically significant at $\alpha = 0.05$ and red for statistically significant at $\alpha = 0.00333$ (from the Bonferroni Correction over 15 tests) when comparing to Quartile 1 of Prop. Black. across the case quartiles and overall

Figures

Figure 1



or

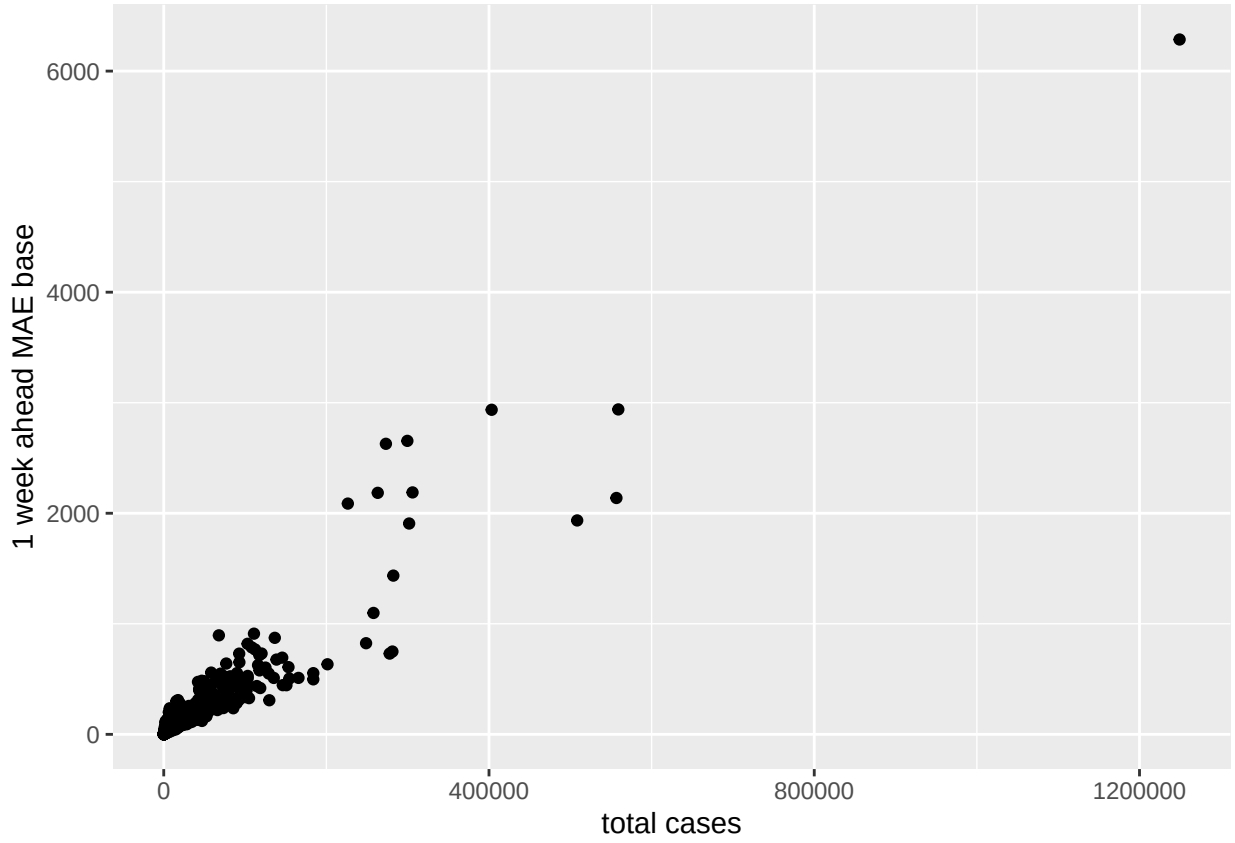

6000
4000
2000
0
1 week ahead MAE base
0
400000
800000
1200000
total cases

Figure 2

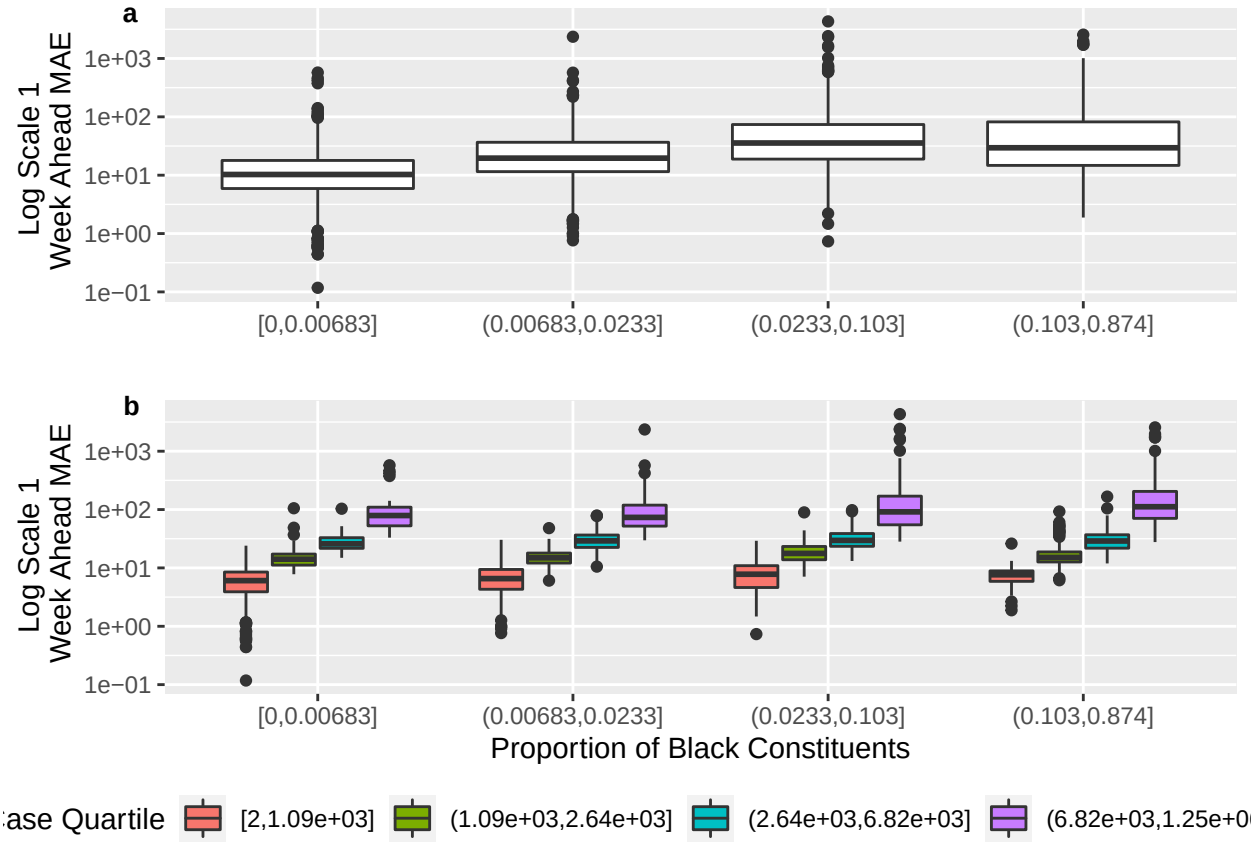


Figure 2 just highest case quartile

```
all_the_data_pull_cutoff %>%
  filter(`Case Quartile`==4) %>%
  ggplot()+
  geom_boxplot(aes(x=as.factor(`black_prop_quart`),
      y=`total cases`)) +
  xlab("Race Quartile amongst Case Quartile 4") +
  ylab("Log Scale Cases") +
  scale_y_log10()
```

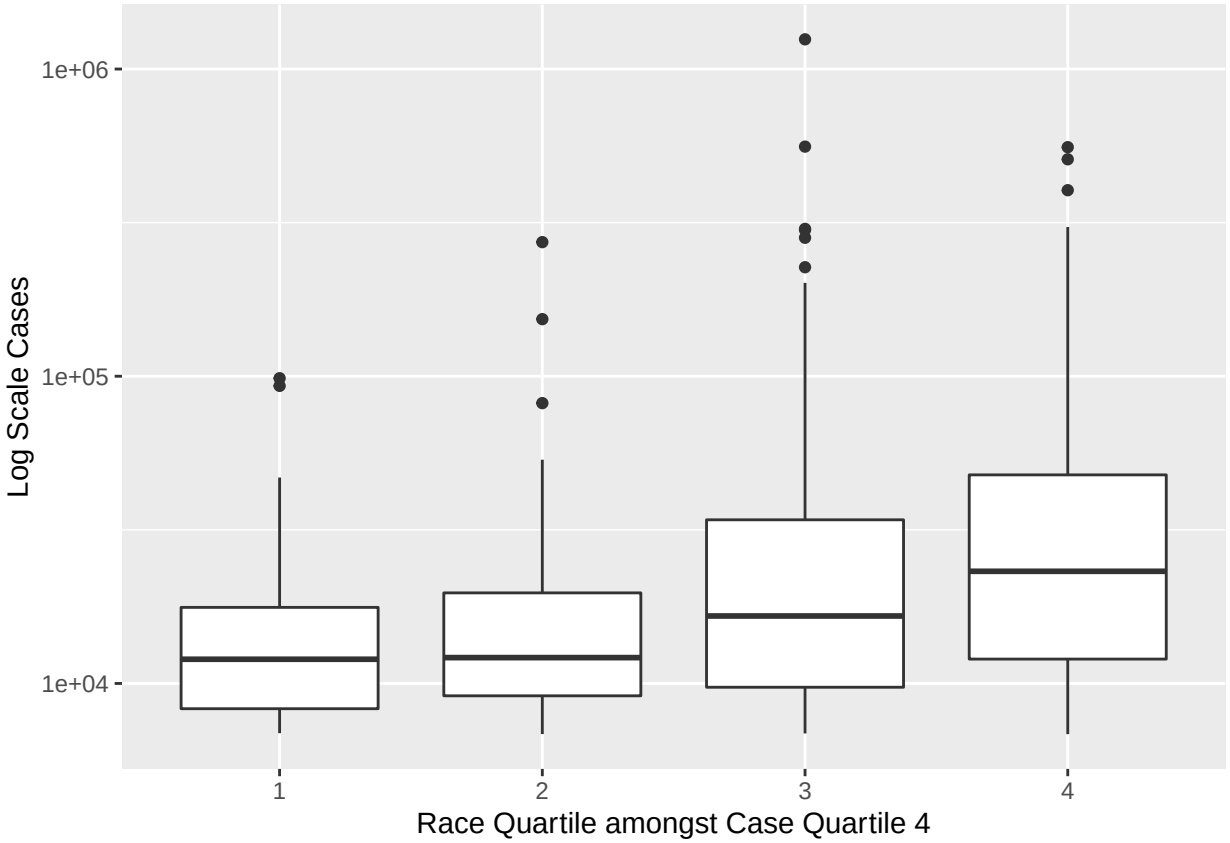

1e+06
1e+05
1e+04
Log Scale Cases
1
2
3
4
Race Quartile amongst Case Quartile 4

Figure 3: 1 and 4 week MAE and RMAE for Racial Demographics

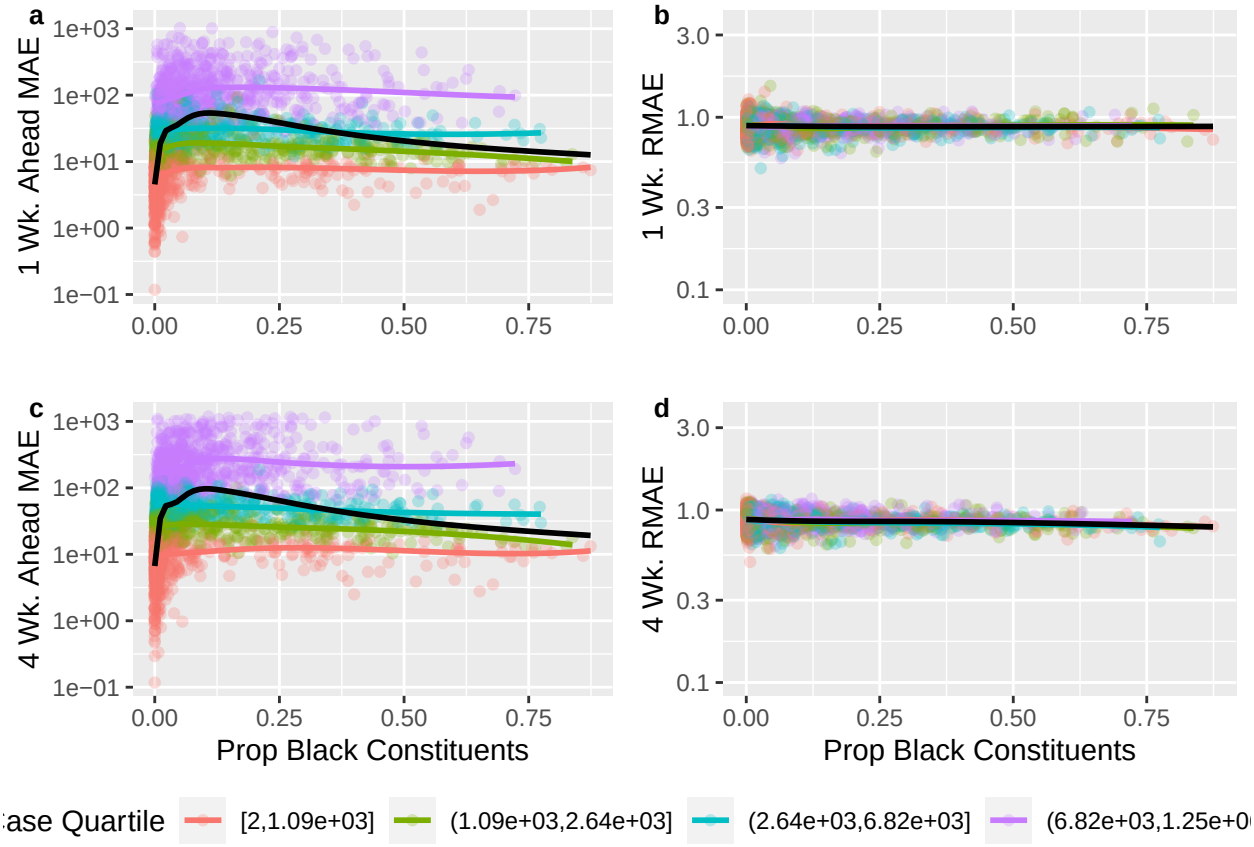


Figure 3 option 2: 1 and 4 week MAE and RMAE for Racial Demographics

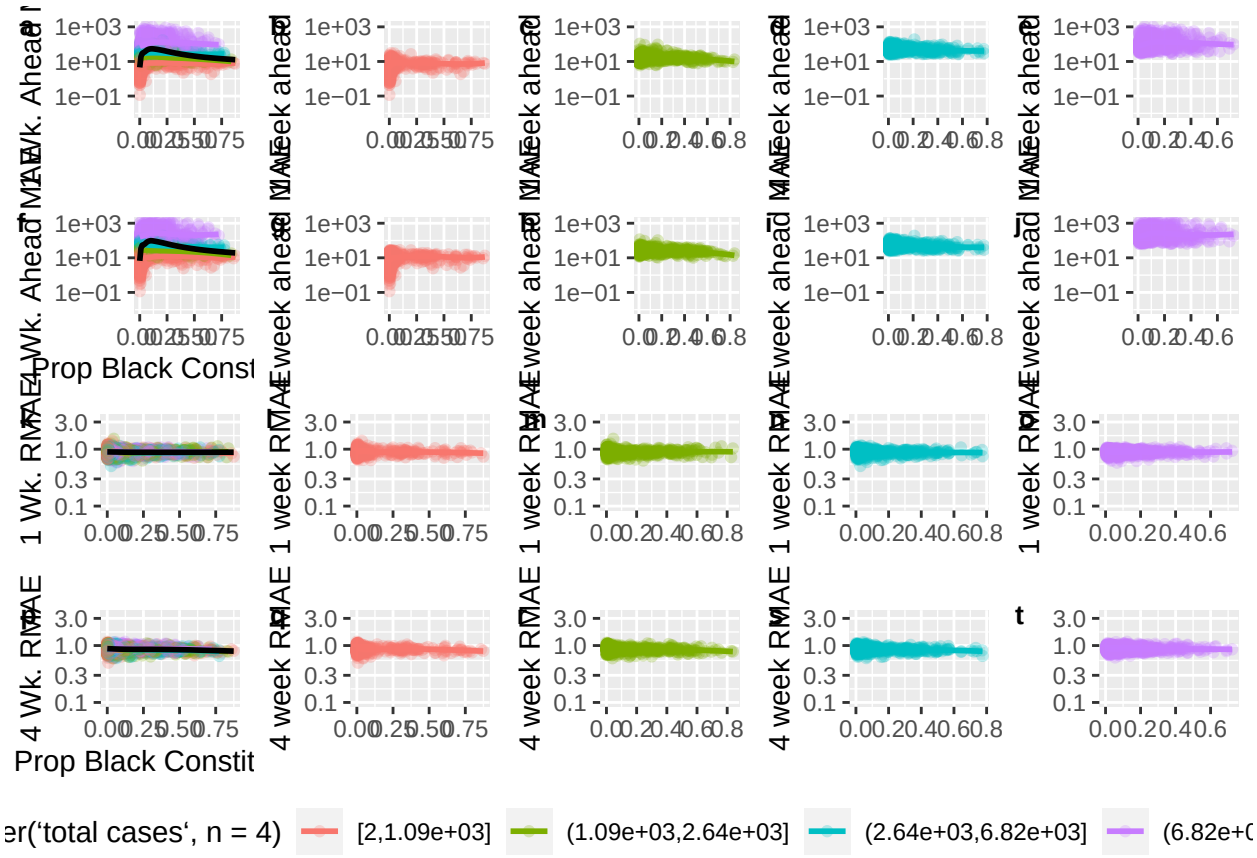
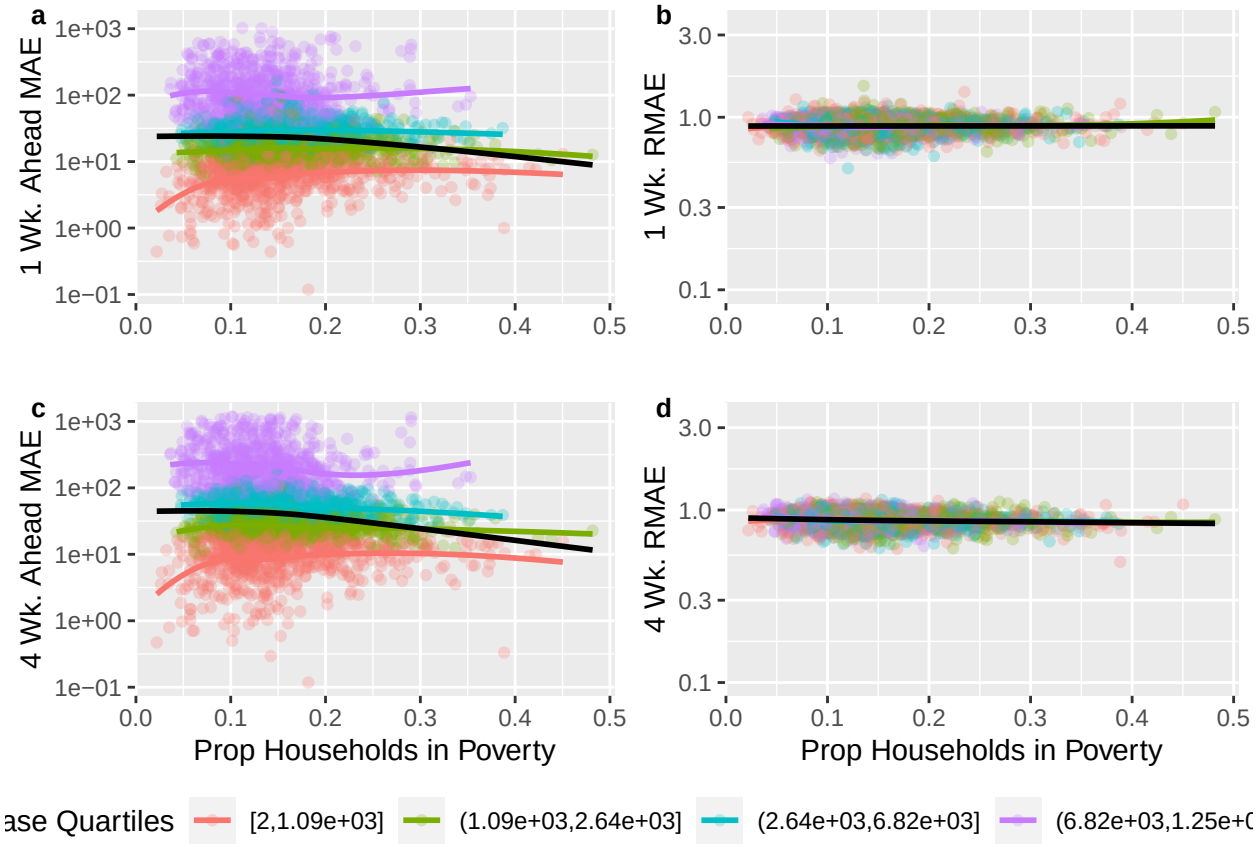


Figure 4: 1 and 4 week MAE and RMAE for Poverty



Tables

Table 1

	R squared
1 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0309
1 wk Ahead MAE as a funct. of Total Population	0.8712
1 wk Ahead MAE as a funct. of Total Cases	0.8862
1 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.8863
4 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0239
4 wk Ahead MAE as a funct. of Total Population	0.9108
4 wk Ahead MAE as a funct. of Total Cases	0.9543
4 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.9554

	R Squared	Adj. R Squared
1 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0309	0.0300
1 wk Ahead MAE as a funct. of Total Population	0.8712	0.8711
1 wk Ahead MAE as a funct. of Total Cases	0.8862	0.8862
1 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.8863	0.8861
4 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0239	0.0230

	R Squared	Adj. R Squared
4 wk Ahead MAE as a funct. of Total Population	0.9108	0.9108
4 wk Ahead MAE as a funct. of Total Cases	0.9543	0.9543
4 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.9554	0.9553

Regression Equations

```
summary(lm(One_Week_MAE~Tot_Cases,data = regressionData))
```

```
##
## Call:
## lm(formula = One_Week_MAE ~ Tot_Cases, data = regressionData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -752.72   -9.19   -5.67    1.31 1238.92
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.330e+01  1.015e+00   13.11   <2e-16 ***
## Tot_Cases   4.046e-03  2.598e-05  155.74   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 54.53 on 3115 degrees of freedom
## Multiple R-squared:  0.8862, Adjusted R-squared:  0.8862
## F-statistic: 2.425e+04 on 1 and 3115 DF,  p-value: < 2.2e-16
```

```
summary(lm(One_Week_MAE~Tot_Cases+Prop_Black*Prop_Pov,
           data = regressionData))
```

```
##
## Call:
## lm(formula = One_Week_MAE ~ Tot_Cases + Prop_Black * Prop_Pov,
##     data = regressionData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -754.78   -9.39   -5.58    1.32 1238.72
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.128e+01  3.112e+00   3.623 0.000296 ***
## Tot_Cases            4.048e-03  2.646e-05 152.979  < 2e-16 ***
## Prop_Black           6.610e+00  2.136e+01   0.309 0.757008
## Prop_Pov             1.843e+01  2.003e+01   0.920 0.357554
## Prop_Black:Prop_Pov -7.790e+01  8.966e+01  -0.869 0.384995
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 54.54 on 3111 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.8863, Adjusted R-squared:  0.8861
## F-statistic:  6062 on 4 and 3111 DF,  p-value: < 2.2e-16
```