Mixed Effect Modeling

Goal of Study

This study examines how different homelessness intervention programs impact the rate of unsheltered people per 100,000.

What is a Mixed-Effects Model?

In simple terms, a mixed-effects model helps us study how different factors influence an outcome while taking into account any variations caused by broader groups or categories (like cities or schools). Mixed-effects models are especially helpful when you're working with data collected across different groups, as they allow us to measure both the overall effect of different programs (like food assistance or emergency shelter) and the unique differences specific to each group (such as city-specific factors that affect homelessness).

Think of it as looking at both the "big picture" effects (which apply generally) and the "smaller picture" effects (specific to each city) all in one model.

Key Terminology

Fixed Effects

- **Explanation**: Fixed effects are the variables we're most interested in studying. These represent the main programs or interventions that we believe have a consistent influence on the outcome across all groups.
- **Example in Context**: In this study, the fixed effects are different homelessness programs (like housing assistance or outreach services). Each program's effect is assumed to be the same across all cities, though the size of the effect can vary based on the program itself.

Random Effects

- **Explanation**: Random effects account for variations that are specific to different groups or contexts. In a mixed-effects model, random effects let us account for how individual groups might influence the outcome differently. Rather than estimating one specific effect per group, the model assumes the influence of each group has a random variation that centers around an average effect.
- **Example in Context**: Here, "city" is a random effect because we expect that each city will have its own unique conditions impacting homelessness rates. The model assumes cities have varying impacts on homelessness but does not treat each city as a separate study; it groups them to get an "average city effect" with some variation.

Intercept

- **Explanation**: The intercept in a mixed-effects model represents the average level of the outcome variable (here, homelessness rate) when all predictors (such as the programs) are set to zero.
- **Example in Context**: The intercept could be thought of as the average homelessness rate in cities without any of these specific programs in place (though in practice, most cities will have at least some services).

Load Data

```
# Regular data
homelessness_data = read.csv('../data/processed/pivoted_and_PIT.csv')

#Standardized Data
standardize = function(col){
  col = (col - mean(col)) / sd(col)
  return(col)
}

homelessness_data_standard = homelessness_data |>
  mutate(across(where(is.numeric) & !all_of("Unsheltered.Per.100.000"), standardize))

# Set the directory where images are saved
output_dir <- "C:/Users/harry/SDCTA/SD-Homelessness-Program-Analysis/Data_Analysis/Visuals/Mixed Effects

# Make sure the directory exists
if(!dir.exists(output_dir)) {
  dir.create(output_dir, recursive = TRUE)
}
```

Standardizing is a process where we adjust the scale of numerical variables so that they have a mean of zero and a standard deviation of one. This helps us compare the impact of different variables more easily, as it puts them all on a similar scale.

In this context, we're standardizing most of the program-related variables to ensure they're comparable and to make interpreting their effects on homelessness rates easier. This also helps the model estimate more effectively, especially when variables have different units or ranges (for instance, comparing dollars spent on housing assistance to the number of outreach services provided).

Why Mixed Effects Model

```
# Examine variation in Unsheltered.Per.100.000 by City
ggplot(homelessness_data, aes(x = City, y = Unsheltered.Per.100.000)) +
  geom_boxplot() +
  labs(title = "Distribution of Unsheltered Rate per 100,000 by City") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Distribution of Unsheltered Rate per 100,000 by City
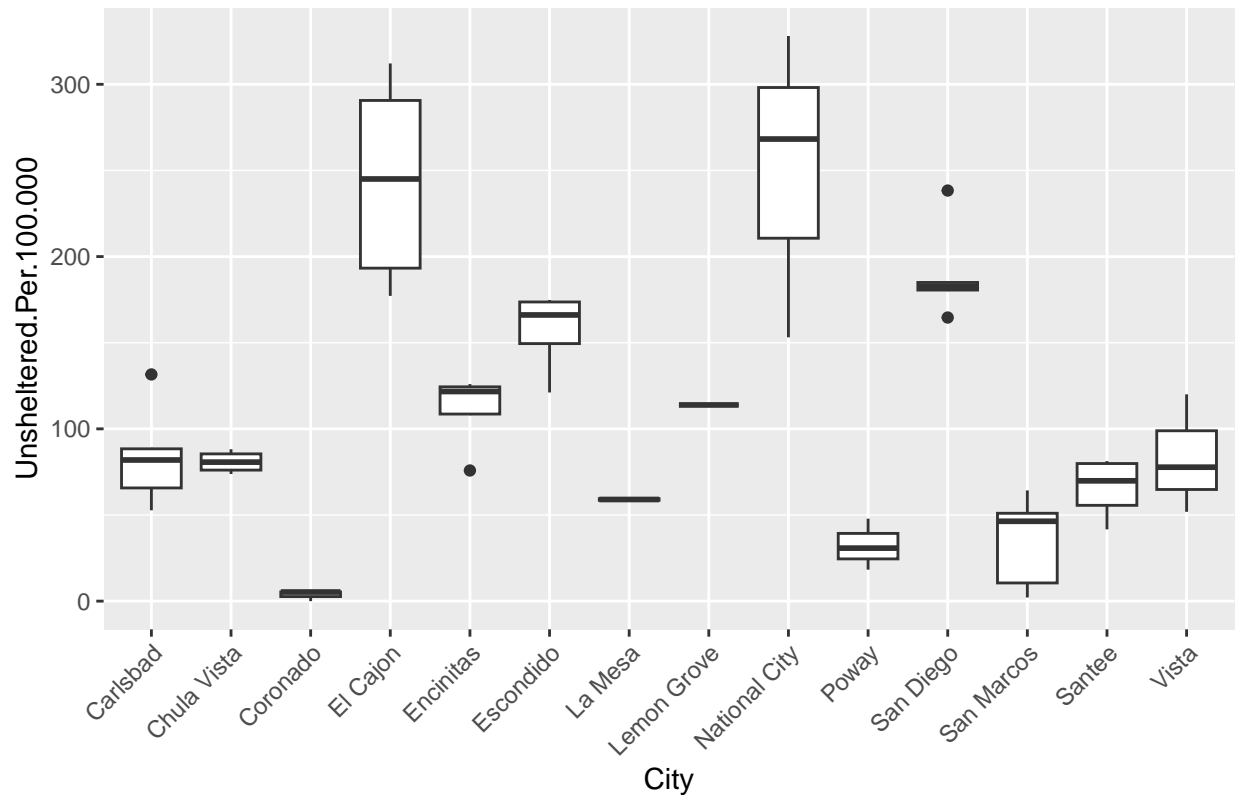

```
ggsave(filename = file.path(output_dir, "City_Unsheltered_Var.png"))
```

```
## Saving 6.5 x 4.5 in image
```

This graph shows how much the unsheltered homelessness rate per 100,000 varies from one city to another. We see that each city has unique conditions affecting homelessness, which suggests that these differences need to be accounted for in our analysis. If we ignored city-level variability, we might mistake city-specific factors as effects of the homelessness programs themselves, which could lead to misleading conclusions.

By using a mixed-effects model, we can include city as a random effect. This allows us to control for city-to-city differences and focus more directly on the main question: how the various homelessness programs affect unsheltered rates overall, while recognizing each city's unique context.

In other words, the mixed-effects model helps us separate the general impact of homelessness programs (fixed effects) from variations that are specific to each city (random effects).

Implement Models

```
ignored_features = c("PEH.Per.100.000", "Total.PEH", "Unsheltered.PEH", "Population")

model <- lmer(Unsheltered.Per.100.000 ~
              bridge.to.housing.network + emergency.shelter + family.reunification.program +
              flexible.funds + food.and.nutrition + homeless.services + homelessness.prevention +
```

```
                homeshare.program + housing.assistance + housing.navigation.services +
                housing.stability.services + motel.voucher + neighborhood.revitalization.services +
                opening.doors.program + outreach + project.h.o.p.e. + rapid.re.housing +
                rental.assistance + restrooms + safe.parking + service.center + staff.and.operations +
                take.back.the.streets + transitional.housing + work.for.hope +
                (1 | City), data = homelessness_data |> select(-all_of(ignored_features)))
```

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```
## Warning: Some predictor variables are on very different scales: consider
## rescaling
## Warning: Some predictor variables are on very different scales: consider
## rescaling
```

```
#Standarized model
modelS <- lmer(Unsheltered.Per.100.000 ~
                bridge.to.housing.network + emergency.shelter + family.reunification.program +
                flexible.funds + food.and.nutrition + homeless.services + homelessness.prevention +
                homeshare.program + housing.assistance + housing.navigation.services +
                housing.stability.services + motel.voucher + neighborhood.revitalization.services +
                opening.doors.program + outreach + project.h.o.p.e. + rapid.re.housing +
                rental.assistance + restrooms + safe.parking + service.center + staff.and.operations +
                take.back.the.streets + transitional.housing + work.for.hope +
                (1 | City), data = homelessness_data_standard |> select(-all_of(ignored_features)))
```

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

Here, we built two versions of a mixed-effects model to assess how various homelessness programs impact unsheltered rates per 100,000 across different cities, treating "city" as a random effect to capture city-specific variations. The first model uses the raw data, while the second uses standardized data to allow for easier comparison of effect sizes among the predictors.

Addressing Warnings

```
vif(model)
```

```
##                                         GVIF Df GVIF^(1/(2*Df))
## bridge.to.housing.network       2.119637e+01  1        4.603952
## emergency.shelter               3.019310e+01  1        5.494825
## family.reunification.program    2.494188e+01  1        4.994185
## flexible.funds                  1.984565e+02  1       14.087458
## food.and.nutrition              1.265858e+00  1        1.125104
## homeless.services               5.727109e+01  1        7.567767
## homelessness.prevention         1.368564e+00  1        1.169856
## homeshare.program               1.279381e+00  1        1.131097
## housing.assistance              6.251358e+00  1        2.500272
## housing.navigation.services     3.900159e+01  1        6.245125
## housing.stability.services      1.993901e+02  1       14.120556
## motel.voucher                   2.016420e+01  1        4.490456
```

```
## neighborhood.revitalization.services 1.312766e+00  1        1.145760
## opening.doors.program                3.907793e+00  1        1.976814
## outreach                             8.250283e+01  1        9.083107
## project.h.o.p.e.                     3.118510e+00  1        1.765930
## rapid.re.housing                     4.773277e+01  1        6.908891
## rental.assistance                    1.401692e+00  1        1.183931
## restrooms                            5.898827e+01  1        7.680382
## safe.parking                         6.839702e+01  1        8.270249
## service.center                       3.003647e+01  1        5.480554
## staff.and.operations                 1.907766e+00  1        1.381219
## take.back.the.streets                2.182301e+00  1        1.477261
## transitional.housing                 2.416204e+01  1        4.915490
## work.for.hope                        3.644808e+13  0            Inf
```

Here, we're checking for multicollinearity in the model using the Variance Inflation Factor (VIF). Multi-collinearity occurs when two or more predictors are highly correlated, making it difficult to isolate the effect of each one. VIF quantifies this by measuring how much the variance of a predictor's estimated coefficient increases due to correlations with other predictors.

A high VIF (above 10, for example) indicates a high level of multicollinearity, while an extremely high VIF (like 4.08e+13 here) shows that the predictor, `work.for.hope`, is almost perfectly collinear with others. This severe multicollinearity causes a "rank deficient" warning, meaning that the model can't estimate unique effects for all predictors. As a result, R automatically removes `work.for.hope` to allow the model to proceed without redundancy.

The warning regarding different scales for predictor variables is only relevant to the raw data, the standardized data addresses this issue.

Visualizing City-Level Random Effects

```
# Random effects for City
ranef_city <- ranef(model)$City
ggplot(ranef_city, aes(x = reorder(row.names(ranef_city), `(Intercept)`), y = `(Intercept)`)) +
  geom_point() +
  coord_flip() +
  labs(title = "City-Level Random Effects", x = "City", y = "Random Intercept Estimate")
```

City–Level Random Effects
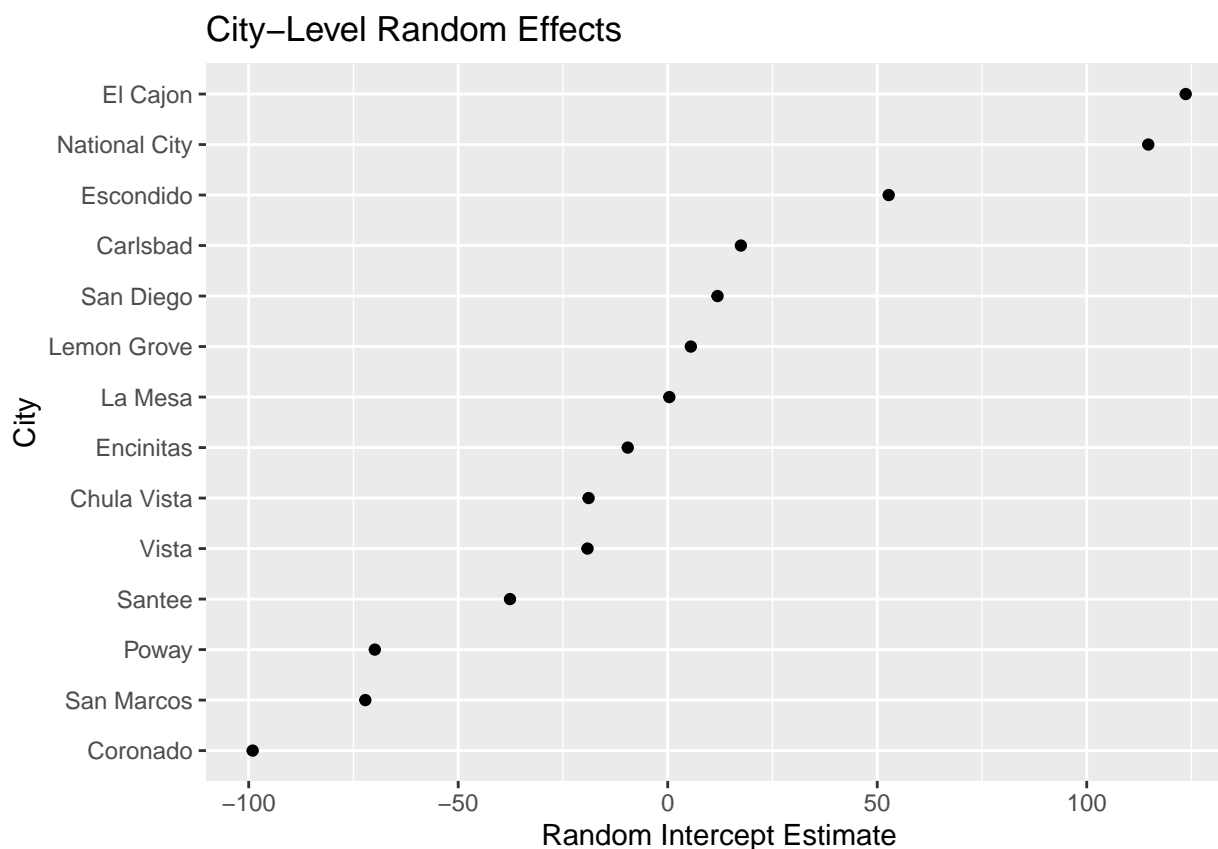

```r
ggsave(filename = file.path(output_dir, "City_Random_Effects.png"))
```

```
## Saving 6.5 x 4.5 in image
```

Each point in the plot represents the random intercept estimate for a city, showing how much the average unsheltered homelessness rate in that city differs from the overall mean rate. Cities like Coronado and San Marcos, for example, have lower-than-average intercepts, suggesting that, after controlling for the program effects, these cities tend to have lower homelessness rates than the average across all cities. Conversely, cities like El Cajon and National City have positive intercept estimates, meaning their rates are generally higher than the overall average. This variation supports the use of a mixed-effects model, as it indicates that city-specific factors significantly impact the outcome and should be accounted for to avoid biased estimates of the program effects.

Model Summaries

```r
summarize_model <- function(model) {
  model_summary <- summary(model)

  # Convert the fixed effects table to a data frame
  fixed_effects <- as.data.frame(model_summary$coefficients)

  # Sort the fixed effects by Estimate
```

	Estimates	Std. Error	df	t value	Pr(>|t|)
bridge.to.housing.network	-0.0041	0.0050058	21.39161	-0.8184315	0.4221347
project.h.o.p.e.	-0.0036	0.0069981	22.92499	-0.5093936	0.6153450
flexible.funds	-0.0019	0.0021488	21.57902	-0.8732437	0.3921444
homeshare.program	-0.0013	0.0016799	23.74414	-0.7459848	0.4629987
neighborhood.revitalization.services	-0.0005	0.0004868	27.31544	-1.0885173	0.2858779
service.center	-0.0002	0.0002378	28.61813	-0.9076437	0.3716465
family.reunification.program	-0.0002	0.0002574	30.27835	-0.8201204	0.4185572
take.back.the.streets	-0.0001	0.0003659	21.93376	-0.3378178	0.7387124
homeless.services	-0.0001	0.0008656	23.65704	-0.1426506	0.8877742
rapid.re.housing	-0.0001	0.0001161	29.49783	-1.0324982	0.3102303
outreach	-0.0001	0.0001333	25.03905	-0.6570337	0.5171535
housing.navigation.services	-0.0001	0.0001576	15.10570	-0.5389667	0.5977642
rental.assistance	-0.0000	0.0000452	22.19783	-0.1551795	0.8780822
emergency.shelter	0.0000	0.0000141	30.90048	1.4946802	0.1451410
housing.assistance	0.0000	0.0000365	23.30938	0.6188417	0.5420219
food.and.nutrition	0.0000	0.0003072	22.96635	0.1315442	0.8964904
transitional.housing	0.0001	0.0000590	29.81341	0.8907320	0.3802051
homelessness.prevention	0.0001	0.0003003	25.23164	0.1799794	0.8586064
staff.and.operations	0.0001	0.0001419	24.67271	1.0374071	0.3096070
safe.parking	0.0002	0.0003522	30.53654	0.4395803	0.6633364
opening.doors.program	0.0005	0.0007804	24.59442	0.6363629	0.5304183
motel.voucher	0.0006	0.0012338	24.46260	0.5239744	0.6050118
housing.stability.services	0.0010	0.0012060	21.93836	0.8076378	0.4279697
restrooms	0.0023	0.0030419	23.14728	0.7604965	0.4546353
(Intercept)	108.2278	24.6232458	13.00387	4.3953524	0.0007233

```r
  sorted_fixed_effects <- fixed_effects[order(fixed_effects$Estimate), ]

  # Add the C-like format for estimates, but remove the predictor names in the formatted column
  sorted_fixed_effects$Estimates <- sprintf("%.4f", sorted_fixed_effects$Estimate)

  # Remove the original 'Estimate' column
  sorted_fixed_effects <- sorted_fixed_effects[, !names(sorted_fixed_effects) %in% c("Estimate")]

  # Rearrange the columns to move 'Estimates' to the front
  sorted_fixed_effects <- sorted_fixed_effects[, c("Estimates", setdiff(names(sorted_fixed_effects), "E

  # Print the modified summary with formatted estimates
  sorted_fixed_effects %>%
    kable("latex", booktabs = TRUE) %>%
    kable_styling(latex_options = "scale_down")
}

summarize_model(model)


#for pdf formatting
cat("\\newpage")
```

	Estimates	Std. Error	df	t value	Pr(>|t|)
flexible.funds	-94.7270	108.477125	21.57901	-0.8732437	0.3921444
rapid.re.housing	-48.2851	46.765330	29.49783	-1.0324982	0.3102303
family.reunification.program	-42.3905	51.688103	30.27835	-0.8201204	0.4185572
outreach	-38.3456	58.361741	25.03905	-0.6570337	0.5171535
service.center	-35.7506	39.388335	28.61813	-0.9076437	0.3716465
bridge.to.housing.network	-29.5786	36.140634	21.39161	-0.8184315	0.4221347
housing.navigation.services	-21.8628	40.564354	15.10570	-0.5389667	0.5977642
neighborhood.revitalization.services	-14.0285	12.887736	27.31544	-1.0885173	0.2858779
homeless.services	-7.3345	51.416086	23.65704	-0.1426506	0.8877742
project.h.o.p.e.	-5.9545	11.689483	22.92499	-0.5093936	0.6153450
homeshare.program	-5.8613	7.857138	23.74414	-0.7459848	0.4629987
take.back.the.streets	-3.3034	9.778653	21.93376	-0.3378178	0.7387124
rental.assistance	-1.2498	8.054077	22.19783	-0.1551795	0.8780822
food.and.nutrition	1.0116	7.689997	22.96635	0.1315442	0.8964904
homelessness.prevention	1.3912	7.729939	25.23164	0.1799794	0.8586064
staff.and.operations	9.8517	9.496502	24.67271	1.0374071	0.3096070
housing.assistance	9.9260	16.039611	23.30938	0.6188417	0.5420219
opening.doors.program	13.0016	20.431050	24.59442	0.6363629	0.5304183
motel.voucher	17.0784	32.593932	24.46260	0.5239744	0.6050118
safe.parking	29.6583	67.469696	30.53654	0.4395803	0.6633364
restrooms	39.2690	51.636027	23.14728	0.7604965	0.4546353
transitional.housing	45.1590	50.698709	29.81341	0.8907320	0.3802051
emergency.shelter	74.7368	50.001877	30.90048	1.4946802	0.1451410
housing.stability.services	87.9868	108.943342	21.93836	0.8076378	0.4279697
(Intercept)	110.4264	21.613701	10.48731	5.1090948	0.0003942

```
summarize_model(modelS)
```

Key Takeaways from the Model's Results

Fixed Effects

Each program's influence is represented by a fixed-effect coefficient, which shows the estimated effect size of that intervention on homelessness rates. A negative coefficient suggests that the program is associated with a decrease in unsheltered homelessness rates, while a positive coefficient implies an increase.

Effect Size and Statistical Significance

The model provides an "Estimate" for each intervention, alongside a "p-value," which indicates the likelihood of the observed effect occurring by chance. Lower p-values (typically below 0.05) suggest a statistically significant impact. For instance, programs like **emergency shelter** may show a significant effect.

However, it's important to note that because we're working with a limited amount of data, it can be challenging to achieve statistical significance. This limitation can make it harder for the model to detect small but meaningful effects, which may lead to high p-values even if a program has a real impact. In future studies, additional data could help in more accurately assessing the significance of these effects.
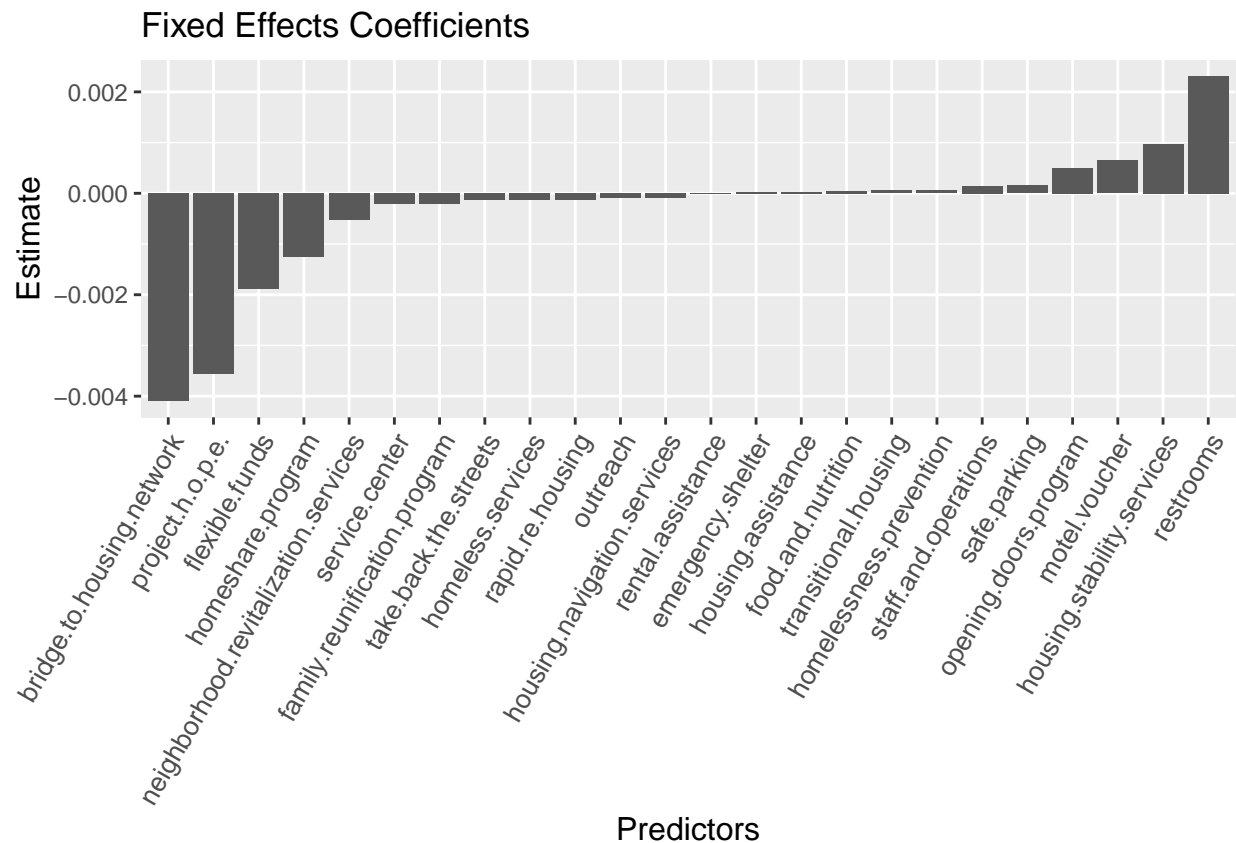
Plot Estimates

```r
# Extract fixed effects
fixed_effects <- as.data.frame(summary(model)$coefficients)

# Remove the intercept row
fixed_effects_no_intercept <- fixed_effects[rownames(fixed_effects) != "(Intercept)", ]

# Reorder the predictors based on the Estimate
fixed_effects_no_intercept$Predictor <- factor(
  rownames(fixed_effects_no_intercept),
  levels = rownames(fixed_effects_no_intercept)[order(fixed_effects_no_intercept$Estimate)]
)

# Plot the fixed effects excluding the intercept, sorted by Estimate
ggplot(fixed_effects_no_intercept, aes(x = Predictor, y = Estimate)) +
  geom_bar(stat = "identity") +
  theme(
    axis.text.x = element_text(angle = 60, hjust = 1, size = 10), # Rotate and adjust the labels
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12)
  ) +
  labs(
    title = "Fixed Effects Coefficients",
    x = "Predictors",
    y = "Estimate"
  )
```

Fixed Effects Coefficients


```r
ggsave(filename = file.path(output_dir, "Estimates_Plot.png"))
```

```
## Saving 6.5 x 4.5 in image
```

```r
# Extract fixed effects
fixed_effectsS <- as.data.frame(summary(modelS)$coefficients)

# Remove the intercept row
fixed_effects_no_interceptS <- fixed_effectsS[rownames(fixed_effectsS) != "(Intercept)", ]

# Reorder the predictors based on the Estimate
fixed_effects_no_interceptS$Predictor <- factor(
  rownames(fixed_effects_no_interceptS),
  levels = rownames(fixed_effects_no_interceptS)[order(fixed_effects_no_interceptS$Estimate)]
)

# Plot the fixed effects excluding the intercept, sorted by Estimate
ggplot(fixed_effects_no_interceptS, aes(x = Predictor, y = Estimate)) +
  geom_bar(stat = "identity") +
  theme(
    axis.text.x = element_text(angle = 60, hjust = 1, size = 10), # Rotate and adjust the labels
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12)
  ) +
  labs(
```
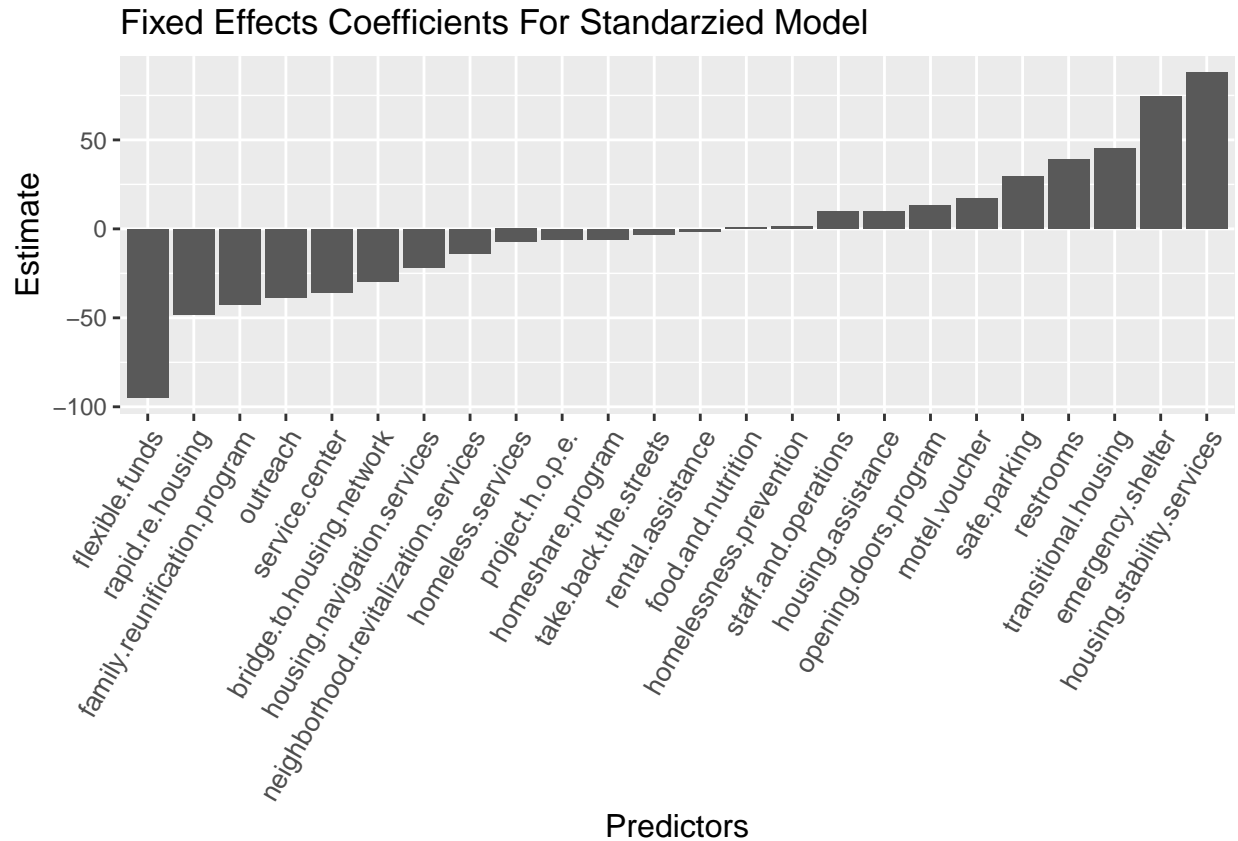
```
    title = "Fixed Effects Coefficients For Standarzied Model",
    x = "Predictors",
    y = "Estimate"
)
```

Fixed Effects Coefficients For Standarzied Model


```
ggsave(filename = file.path(output_dir, "Estimates_Plot_Standardized.png"))
```

```
## Saving 6.5 x 4.5 in image
```

Plot Estimates

The plot of fixed-effect estimates helps us visualize the relative influence of each intervention on homelessness rates. Each bar represents a different program, with the height indicating the estimated effect size (or "Estimate") for that intervention.

Interpreting the Plot

- **Positive Estimates**: Programs with positive bars are associated with an increase in homelessness rates, according to the model. This doesn't necessarily mean these programs cause homelessness to rise; it could be that they are implemented in areas with greater homelessness needs, or that other factors influencing homelessness were not accounted for in the model.
- **Negative Estimates**: Programs with negative bars are associated with a reduction in unsheltered homelessness rates. These programs might be more effective at reducing homelessness, or they may operate in areas where homelessness rates are naturally lower.
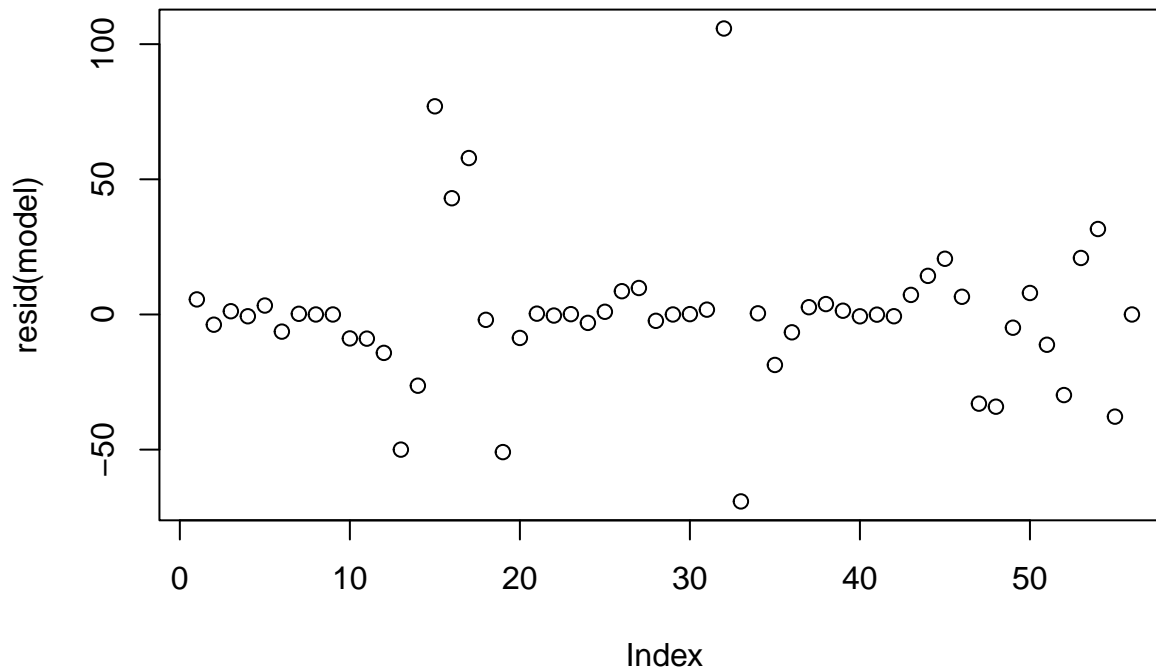
Assessing Residuals and Model Fit

```r
png(filename = file.path(output_dir, "Residual_Plot.png"))
plot(resid(model))
dev.off()
```

```
## pdf
##   2
```

```r
png(filename = file.path(output_dir, "QQ_Plot.png"))
qqnorm(resid(model))
qqline(resid(model))
dev.off()
```
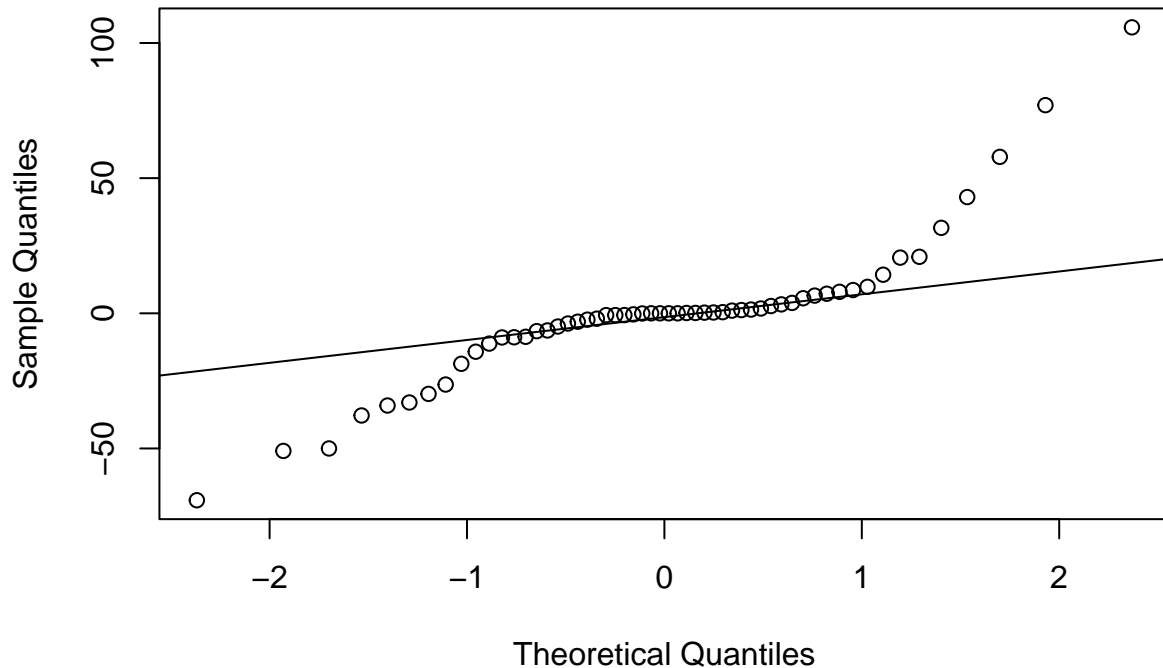
```
## pdf
##   2
```

```r
plot(resid(model))
```



```r
qqnorm(resid(model))
qqline(resid(model))
```

Normal Q–Q Plot

Evaluating residuals and model fit helps determine how well the model represents the data and whether its assumptions hold. Residuals are the differences between observed homelessness rates and those predicted by the model. Ideally, residuals should be randomly distributed around zero, indicating that the model's predictions are unbiased.

Plotting Residuals

The residual plot for this model shows that the residuals are fairly evenly scattered around zero, with no obvious patterns, indicating that the model's assumption of **constant variance** (homoscedasticity) is largely met. This suggests that the model estimates are generally reliable and that there isn't a systematic error affecting the predictions.

Normality of Residuals

The Q-Q plot of residuals shows that most points fall along the line, suggesting that the residuals approximately follow a normal distribution. However, a few points deviate from the line at both extremes, indicating the presence of some outliers. While these deviations are not extreme, they do suggest a slight departure from normality, which could impact the model's assumptions in certain cases.

Outliers and Their Influence

The model identified several observations with large residuals, indicating they are potential outliers. Specifically, cities with the highest and lowest residuals differ significantly from the overall trend, which could mean they have unique characteristics or exceptional homelessness rates not fully accounted for by the model.

Results if outliers are removed

```r
# Extract residuals from the model
residuals <- resid(model)

# Set a threshold for identifying outliers
# For example, residuals greater than 2 standard deviations from the mean
threshold <- 2 * sd(residuals)

# Identify the indices of the outliers
outlier_indices <- which(abs(residuals) > threshold)

# Display the indices and residual values of the outliers
outliers <- data.frame(Index = outlier_indices, Residual = residuals[outlier_indices])
print(outliers)
```

```
##    Index  Residual
## 15    15  77.00378
## 17    17  57.86662
## 32    32 105.78546
## 33    33 -69.15055
```

```r
# View the corresponding rows in the original data
outlier_data <- homelessness_data[outlier_indices, ]

# Remove the identified outlier indices from the dataset
homelessness_data_no_outliers <- homelessness_data[-c(15, 17, 32, 33), ]

# Re-run the model without outliers
model_no_outliers <- lmer(Unsheltered.Per.100.000 ~
            bridge.to.housing.network + emergency.shelter + family.reunification.program +
            flexible.funds + food.and.nutrition + homeless.services + homelessness.prevention +
            homeshare.program + housing.assistance + housing.navigation.services +
            housing.stability.services + motel.voucher + neighborhood.revitalization.services +
            opening.doors.program + outreach + project.h.o.p.e. + rapid.re.housing +
            rental.assistance + restrooms + safe.parking + service.center + staff.and.operations +
            take.back.the.streets + transitional.housing + work.for.hope +
            (1 | City),
            data = homelessness_data_no_outliers |> select(-all_of(ignored_features)))
```

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```r
summarize_model(model_no_outliers)
```

```r
#Standardized Model with no outliers
homelessness_data_no_outliers_standard = homelessness_data_no_outliers |>
  mutate(across(where(is.numeric) & !all_of("Unsheltered.Per.100.000"), standardize))

model_no_outliersS <- lmer(Unsheltered.Per.100.000 ~
            bridge.to.housing.network + emergency.shelter + family.reunification.program +
```

	Estimates	Std. Error	df	t value	Pr(>|t|)
bridge.to.housing.network	-0.0041	0.0032795	16.37769	-1.2521319	0.2281088
project.h.o.p.e.	-0.0040	0.0046361	18.31420	-0.8684920	0.3963638
homeshare.program	-0.0020	0.0012874	24.76485	-1.5869303	0.1252158
flexible.funds	-0.0019	0.0014165	16.82570	-1.3721891	0.1880227
neighborhood.revitalization.services	-0.0007	0.0003552	24.30181	-1.9914134	0.0577899
service.center	-0.0003	0.0001642	24.83715	-1.6688620	0.1077038
family.reunification.program	-0.0002	0.0001873	25.04342	-1.1082740	0.2782742
take.back.the.streets	-0.0001	0.0002430	17.46802	-0.5006767	0.6228484
rapid.re.housing	-0.0001	0.0000823	26.51813	-1.3802994	0.1790203
homeless.services	-0.0001	0.0005910	19.52383	-0.1849672	0.8551627
housing.navigation.services	-0.0001	0.0001261	12.07971	-0.5688731	0.5798655
outreach	-0.0000	0.0000920	19.77669	-0.4843383	0.6334663
rental.assistance	-0.0000	0.0000315	20.40753	-0.1264918	0.9005806
safe.parking	0.0000	0.0002536	26.88074	0.0525117	0.9585093
emergency.shelter	0.0000	0.0000101	26.88842	2.0604259	0.0491546
housing.assistance	0.0000	0.0000249	18.80925	1.6689511	0.1116873
transitional.housing	0.0001	0.0000419	26.14695	1.7741930	0.0876782
food.and.nutrition	0.0001	0.0002025	17.18965	0.3930064	0.6991459
homelessness.prevention	0.0002	0.0002050	19.65064	0.7372626	0.4696750
staff.and.operations	0.0003	0.0001519	25.55805	1.9618820	0.0607534
opening.doors.program	0.0007	0.0005440	21.41411	1.2784357	0.2147790
motel.voucher	0.0007	0.0008386	20.95343	0.8603854	0.3993147
housing.stability.services	0.0010	0.0007981	17.37370	1.2827368	0.2164342
restrooms	0.0017	0.0020135	17.65967	0.8633347	0.3995273
(Intercept)	95.0640	21.0178262	11.58061	4.5230182	0.0007629

	Estimates	Std. Error	df	t value	Pr(>|t|)
flexible.funds	-101.8162	74.199805	16.825696	-1.3721891	0.1880227
rapid.re.housing	-47.4330	34.364296	26.518128	-1.3802994	0.1790203
service.center	-47.0326	28.182460	24.837153	-1.6688620	0.1077038
family.reunification.program	-43.1701	38.952566	25.043419	-1.1082740	0.2782742
bridge.to.housing.network	-30.7442	24.553502	16.377690	-1.2521319	0.2281088
outreach	-20.2259	41.759773	19.776689	-0.4843383	0.6334663
neighborhood.revitalization.services	-19.3919	9.737759	24.301811	-1.9914134	0.0577899
housing.navigation.services	-19.1628	33.685575	12.079714	-0.5688731	0.5798655
homeshare.program	-9.9161	6.248606	24.764852	-1.5869303	0.1252158
project.h.o.p.e.	-6.9795	8.036401	18.314204	-0.8684920	0.3963638
homeless.services	-6.7294	36.381526	19.523826	-0.1849672	0.8551627
take.back.the.streets	-3.3750	6.740922	17.468016	-0.5006767	0.6228484
rental.assistance	-0.7355	5.814736	20.407533	-0.1264918	0.9005806
food.and.nutrition	2.0652	5.254896	17.189654	0.3930064	0.6991459
safe.parking	2.6423	50.319106	26.880741	0.0525117	0.9585093
homelessness.prevention	4.0358	5.474011	19.650644	0.7372626	0.4696750
opening.doors.program	18.8671	14.757931	21.414111	1.2784357	0.2147790
housing.assistance	18.9042	11.326979	18.809251	1.6689511	0.1116873
motel.voucher	19.7418	22.945296	20.953429	0.8603854	0.3993147
staff.and.operations	20.5038	10.451089	25.558048	1.9618820	0.0607534
restrooms	30.6133	35.459389	17.659670	0.8633347	0.3995273
transitional.housing	66.1363	37.276858	26.146954	1.7741930	0.0876782
emergency.shelter	76.0686	36.918866	26.888424	2.0604259	0.0491546
housing.stability.services	95.9683	74.815254	17.373696	1.2827368	0.2164342
(Intercept)	105.6778	18.042531	9.264697	5.8571506	0.0002159

```
            flexible.funds + food.and.nutrition + homeless.services + homelessness.prevention +
            homeshare.program + housing.assistance + housing.navigation.services +
            housing.stability.services + motel.voucher + neighborhood.revitalization.services +
            opening.doors.program + outreach + project.h.o.p.e. + rapid.re.housing +
            rental.assistance + restrooms + safe.parking + service.center + staff.and.operations +
            take.back.the.streets + transitional.housing + work.for.hope +
            (1 | City),
            data = homelessness_data_no_outliers_standard |> select(-all_of(ignored_features)))
```

fixed-effect model matrix is rank deficient so dropping 1 column / coefficient

```
summarize_model(model_no_outliersS)
```

```
#Graph estmates for models without ouliers
# Extract fixed effects
fixed_effects <- as.data.frame(summary(model)$coefficients)

# Remove the intercept row
fixed_effects_no_intercept <- fixed_effects[rownames(fixed_effects) != "(Intercept)", ]

# Reorder the predictors based on the Estimate
fixed_effects_no_intercept$Predictor <- factor(
```
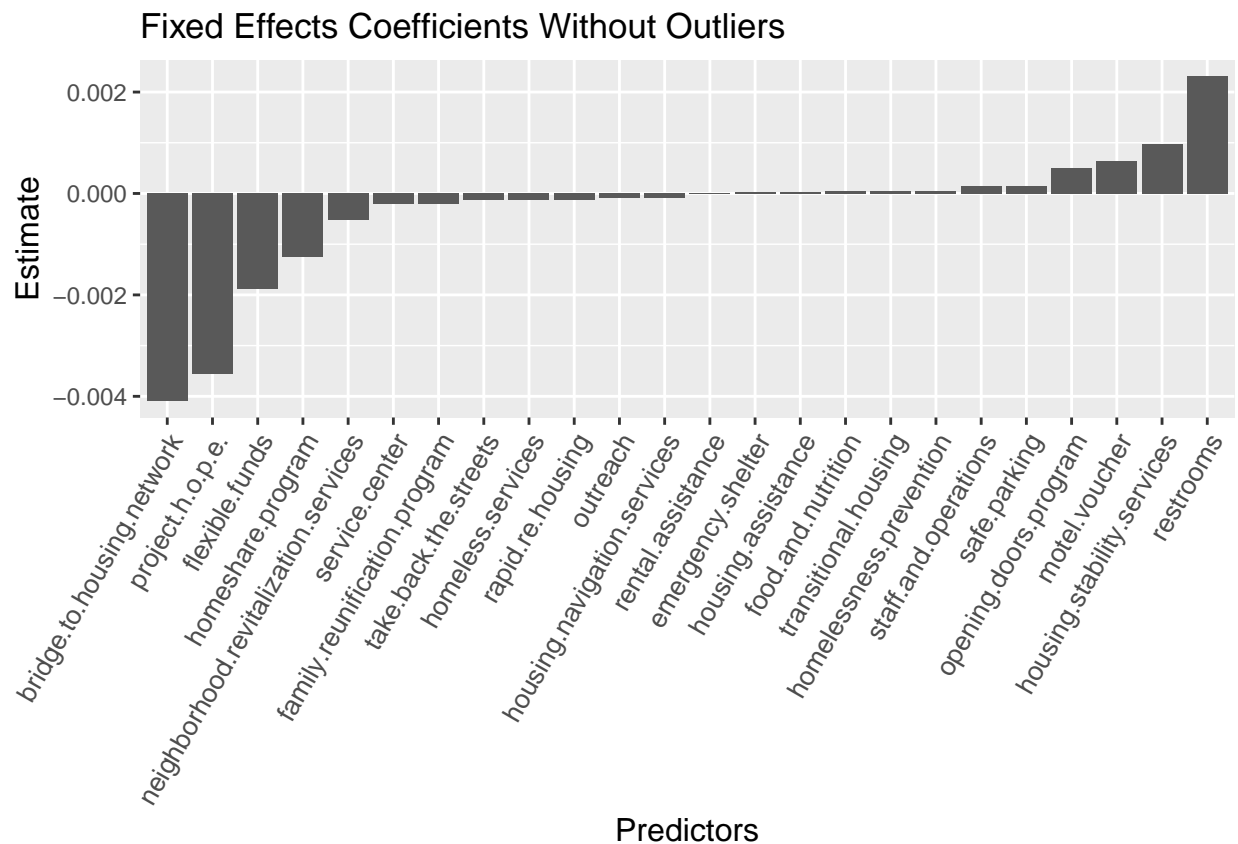
```
    rownames(fixed_effects_no_intercept),
    levels = rownames(fixed_effects_no_intercept)[order(fixed_effects_no_intercept$Estimate)]
)

# Plot the fixed effects excluding the intercept, sorted by Estimate
ggplot(fixed_effects_no_intercept, aes(x = Predictor, y = Estimate)) +
  geom_bar(stat = "identity") +
  theme(
    axis.text.x = element_text(angle = 60, hjust = 1, size = 10), # Rotate and adjust the labels
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12)
  ) +
  labs(
    title = "Fixed Effects Coefficients Without Outliers",
    x = "Predictors",
    y = "Estimate"
  )
```


Fixed Effects Coefficients Without Outliers

```
ggsave(filename = file.path(output_dir, "Estimates_Plot_No_Outliers.png"))
```

```
## Saving 6.5 x 4.5 in image
```

```
# Extract fixed effects
fixed_effectsS <- as.data.frame(summary(model_no_outliersS)$coefficients)
```
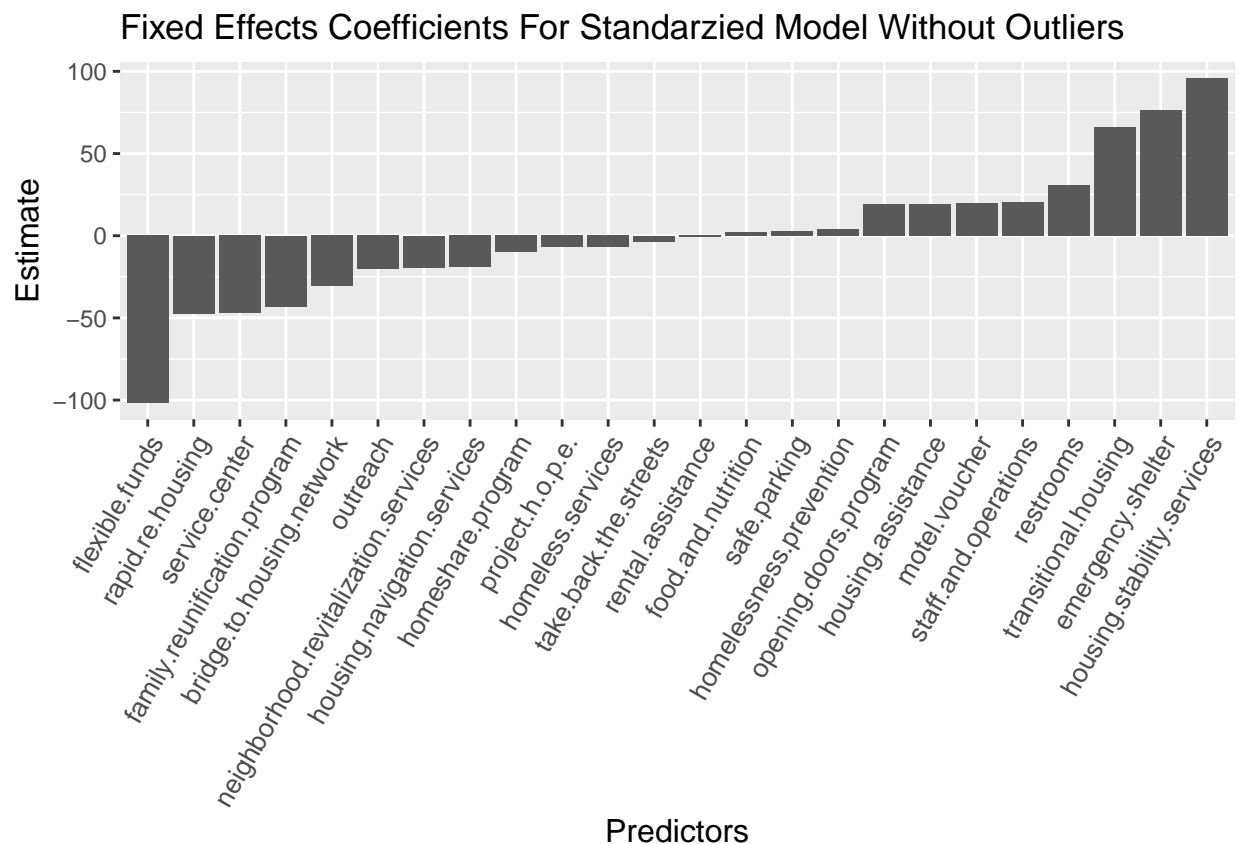
17

```
# Remove the intercept row
fixed_effects_no_interceptS <- fixed_effectsS[rownames(fixed_effectsS) != "(Intercept)", ]

# Reorder the predictors based on the Estimate
fixed_effects_no_interceptS$Predictor <- factor(
  rownames(fixed_effects_no_interceptS),
  levels = rownames(fixed_effects_no_interceptS)[order(fixed_effects_no_interceptS$Estimate)]
)

# Plot the fixed effects excluding the intercept, sorted by Estimate
ggplot(fixed_effects_no_interceptS, aes(x = Predictor, y = Estimate)) +
  geom_bar(stat = "identity") +
  theme(
    axis.text.x = element_text(angle = 60, hjust = 1, size = 10), # Rotate and adjust the labels
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12)
  ) +
  labs(
    title = "Fixed Effects Coefficients For Standarzied Model Without Outliers",
    x = "Predictors",
    y = "Estimate"
  )
```


Fixed Effects Coefficients For Standarzied Model Without Outliers

```r
ggsave(filename = file.path(output_dir, "Estimates_Plot_No_Outliers_Standardized.png"))
```

```
## Saving 6.5 x 4.5 in image
```

```r
# Residuals and Normality without outliers
png(filename = file.path(output_dir, "Residual_Plot_No_Outliers.png"))
plot(resid(model_no_outliers))
dev.off()
```
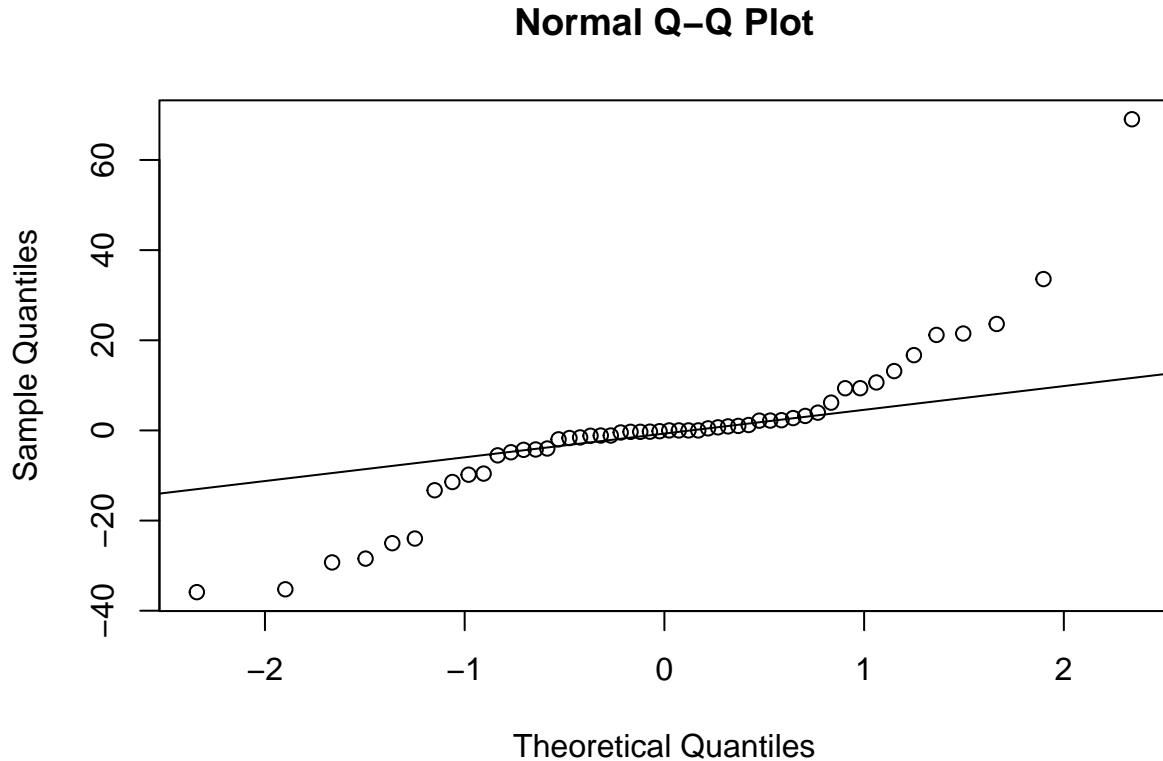
```
## pdf
##   2
```

```r
png(filename = file.path(output_dir, "QQ_Plot_No_Outliers.png"))
qqnorm(resid(model_no_outliers))
qqline(resid(model_no_outliers))
dev.off()
```

```
## pdf
##   2
```

```r
plot(resid(model_no_outliers))
```

```
qqnorm(resid(model_no_outliers))
qqline(resid(model_no_outliers))
```

Normal Q–Q Plot
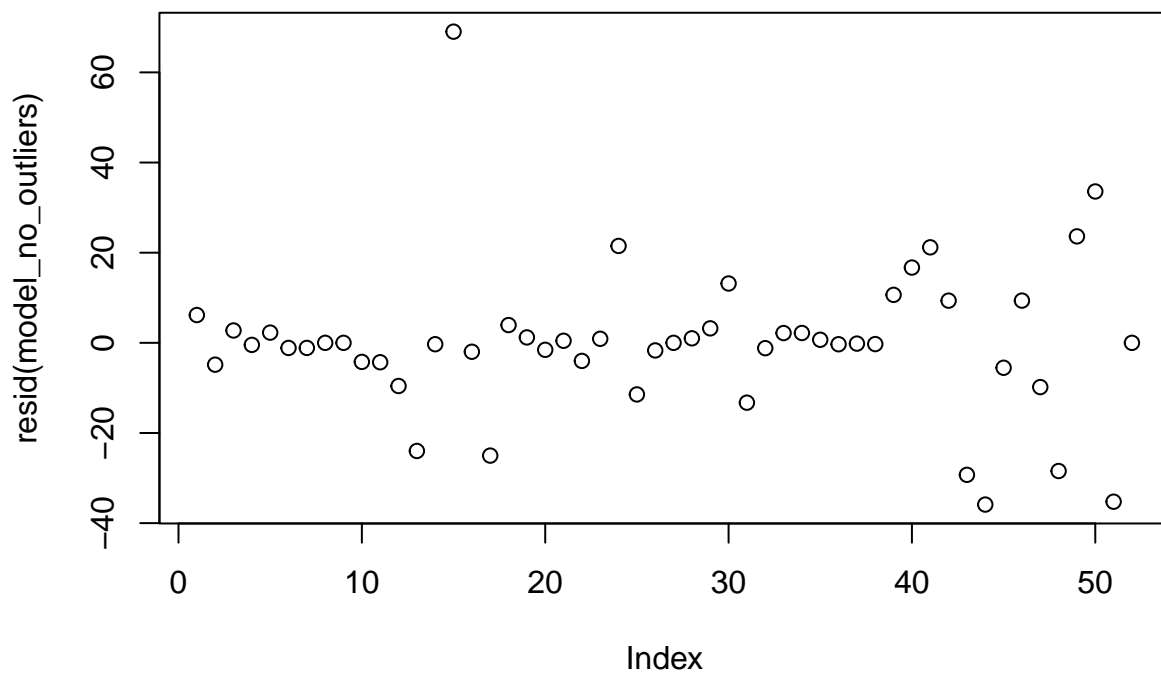
The order of most to least efficient programs remains consistent even when outliers are removed, only the extent to which they affect Unsheltered PEH per 100,000 is affected, though not by much.

Conclusion and Results

This analysis assessed the effectiveness of various homelessness intervention programs across different cities, using a mixed-effects model to account for both program impacts (fixed effects) and city-specific influences (random effects). The model allowed us to estimate the general effect of each intervention while recognizing variations between cities, providing an understanding of how different programs contribute to homelessness rates.

Key Findings

- **Program Impacts**: The fixed effects estimates reveal that some programs are associated with increases in unsheltered homelessness rates, while others show potential for reductions.
- **Top 3 Programs Associated with Increased Homelessness Rates**:
  - **Housing stability services** and **emergency shelter** are among the highest, indicating a strong association with increased homelessness rates. This suggests that these services might be implemented in areas with higher homelessness needs or reflect other unmeasured factors.
  - **Transitional housing** also has a high positive estimate, which could mean it is commonly deployed in areas with more severe homelessness.

- **Top 3 Programs Associated with Decreased Homelessness Rates**:
    - **Flexible funds** shows the largest negative effect, suggesting it is associated with reductions in homelessness rates. This program may provide immediate, adaptable support that helps prevent or mitigate homelessness.
    - **Rapid rehousing** and **family reunification** also have strong negative estimates, highlighting their potential impact in reducing unsheltered homelessness rates.

Model Fit and Reliability

Residual analysis shows that the model fits the data reasonably well, with residuals displaying no major patterns and an approximate normal distribution. The presence of some outliers does indicate that certain cities have unique homelessness rates not fully explained by the model, but these do not significantly alter the overall results. This suggests that the model's conclusions are robust and reliable within the limits of the data available.

Limitations and Future Directions

While the model provides valuable insights, it is based on a limited dataset, which affects the statistical power to detect significant effects for all interventions. Some programs with high p-values may still have meaningful impacts that are hard to capture with the current data. Additional data could improve model precision, help validate these findings, and allow for a more thorough understanding of each program's impact.

In summary, this report highlights the complex nature of homelessness interventions and the need for targeted approaches based on local contexts. The analysis offers a foundation for understanding which types of programs may be more effective and underscores the potential for further research to refine these findings. With continued data collection and more nuanced models, policymakers can better allocate resources to the interventions that are likely to make the most impact on reducing homelessness.