Assessing Mental Health Outcomes Across LGB Individuals from Generations Survey

Shirley Mathur

4/14/2022

Introduction

Background

Even in modern American society, lesbian, gay, and bisexual (LGB) individuals are still targets of discrimination at a systemic level, as evidenced by the legislation that seeks to target these individuals at all levels of government, ranging from the federal to municipal ("Laws &Amp; Legislation - Hrc," n.d.). As posited by Meyer, LGB individuals that have to face such discrimination and prejudice in society are subject to an increased risk of developing mental health problems due to the *minority stress* that they encounter in their daily lives Meyer (2003). However, while this original meta-analysis specifically explored the disparities between the mental health outcomes of LGB individuals against heterosexual individuals, it did not necessarily account for compounding factors that could also influence stress, such as a LGB individuals who are also part of racial minority groups or comparing LGB individuals who are women as opposed to LGB individuals who are men. More recent research exploring these same phenomenon has been done in recent years to see how over time the mental health outcomes of LGB individuals compare to those of heterosexual individuals, while also trying to adjust for these other factors that account for how LGB individuals might also be part of other minority groups. One such study, utilizing data from the National Epidemiologic Survey of Alcohol and Related Conditions (NESARC) that was collected in 2005, found that LGB individuals were associated with a higher odds of having mental health disorders only if they were subject to other forms of discrimination as due to their other marginalized identities outside of just being a LGB individual Bostwick et al. (2014).

In this report, I seek to further expand upon the line of exploration of how different marginalized identities could together have an association with higher odds of having mental health disorders. I also seek to control for other factors that might influence differential treatment for individuals based on their societal contexts, such as accounting for the region in which people live.

Dataset Description and Variables of Interest

The dataset that was used for this report consisted of survey data from Wave 1 and Wave 2 of the Generations study that was conducted from 2017 through 2019. This study sought to explore a variety of outcomes, ranging from personal and social wellbeing to health outcomes, for members of the LGB population in the US. With context to the aims of my investigation, I was mainly interested in demographic variables that described the background characteristics of the surveyed individuals and how those were associated with mental health outcomes. The variables of interest are delineated below.

- Kessler-6 Score - the Kessler-6 is a questionnaire that asks participants how often they have felt "nervous," "hopeless," "restless or fidgety," "so depressed that nothing could cheer you up," "that everything was an effort," and "worthless" in the past 30 days, and is used as a clinical screening tool for assessing if an individual is suffering from Major Depressive Disorder (MDD) Kessler et al. (2003). This score was the outcome of interest in assessing the mental health outcomes of respondents.
- Sexuality - this variable is a categorical variables that categorizes respondents as either "gay/lesbian", "bisexual", or "other", depending on the write-in responses they provided for how they label their

sexuality.

- Education - this variable is a categorical variables that provides information on a respondent's level of education as "high school or less", "some college", "college completed", or "more than college completed."
- Urbanicity - this variable encodes whether a respondent lives in a locale that is considered urban, as defined in the Generations study technical notes, or if a respondent lives in a locale that is not considered urban.
- Geographical Census Region - encodes which geographical census region a respondent is from, and the census regions are defined as
    - Northeast: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
    - Midwest: Indiana, Illinois, Michigan, Ohio, Wisconsin, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
    - South: Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, Texas
    - West: Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming, Alaska, California, Hawaii, Oregon, Washington
- Cohort - variable that characterizes whether a respondent belongs to the "young" (16-27), "middle" (32-43), and "old" (50-61) age range.
- Race - variable that characterizes the self-reported race of a respondent as being one of: White, Hispanic/Latino, Asian, Multiracial, Middle Eastern, American Indian, Black/African American, or Native Hawaiian/Pacific Islander.
- Sex and Gender - variable that combines sex and gender information to classify respondents as one of: women, non-transgender; men, non- transgender; genderqueer/non- binary (GQNB), female; GQNB, male.
- Poverty - variable that characterizes whether a respondent lives in poverty or not, where living in poverty is defined as living below 100% Federal Poverty Line (FPL) according to the FPL calculated in 2017.

For the final dataset, I selected the above variables from the original dataset that contained 894 observations for participants who participated in both Wave 1 and Wave 2 and 929 variables, of which majority were survey questions that were used to compute measures for outcomes of interest (such as the Kessler-6 Score). I removed all observations that had missing data for any of the variables of interest, which resulted in dropping 57 observations and a final dataset containing 837 observations.

Objective

The main objective of this report is to explore the associations between various demographic characteristics of a LGB individual and their mental health outcomes. Specifically, in the context of this study, since participants were measured at Wave 1 and Wave 2, I want to examine how, if at all, a participant's mental health outcome changed over the course of the year and what associations this change might have with the background characteristics and an individual. This would allow us to gain insight into how other societal factors and minority statuses that LGB individuals hold might impact their mental health outcomes.

The secondary objective of this report is to explore if LGB individuals across different races had varying mental health outcomes over the course of this study. Thus, I aim to explore whether or not the direction and magnitude of an association with a LGB individual being from a particular racial category with mental health outcomes differs across the racial categories. This question is of particular interest to me as race is one aspect of an individual's identity that is theorized as leading to *minority stress* in an individual's life, so I wish to examine the intersection of these stressors in context with the stressors that are commonly felt amongst LGB individuals.

Exploratory Data Analysis

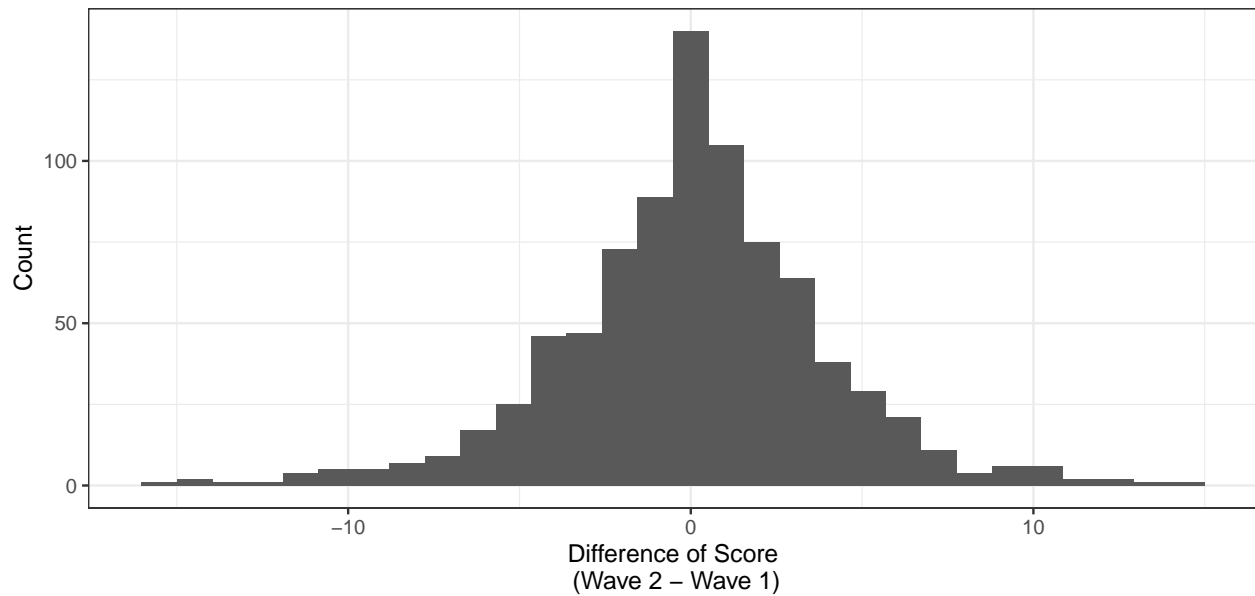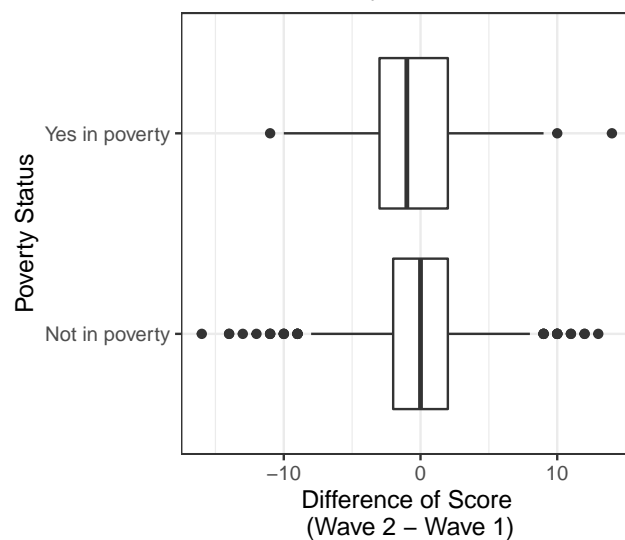### Distribution in Change of Kessler–6 Scores Among Participants

Figure 1: Distribution in change of Kessler–6 Scores from Wave 1 survey to Wave 2 survey. As seen above, the distribution of difference in score appears to be approximately normal.

A

Change of Kessler–6 Scores Across Poverty Status
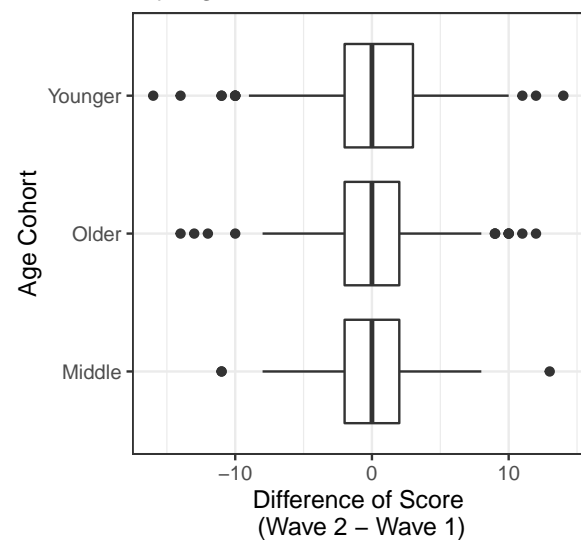
B

Change of Kessler–6 Scores by Age Cohort

Figure 2: Distribution of changes in Kessler–6 scores from Wave 1 survey to Wave 2 survey amongst participants with varying demographic characeristics of poverty status and age cohort.

A   B

Change of Kessler−6 Scores
Across Race

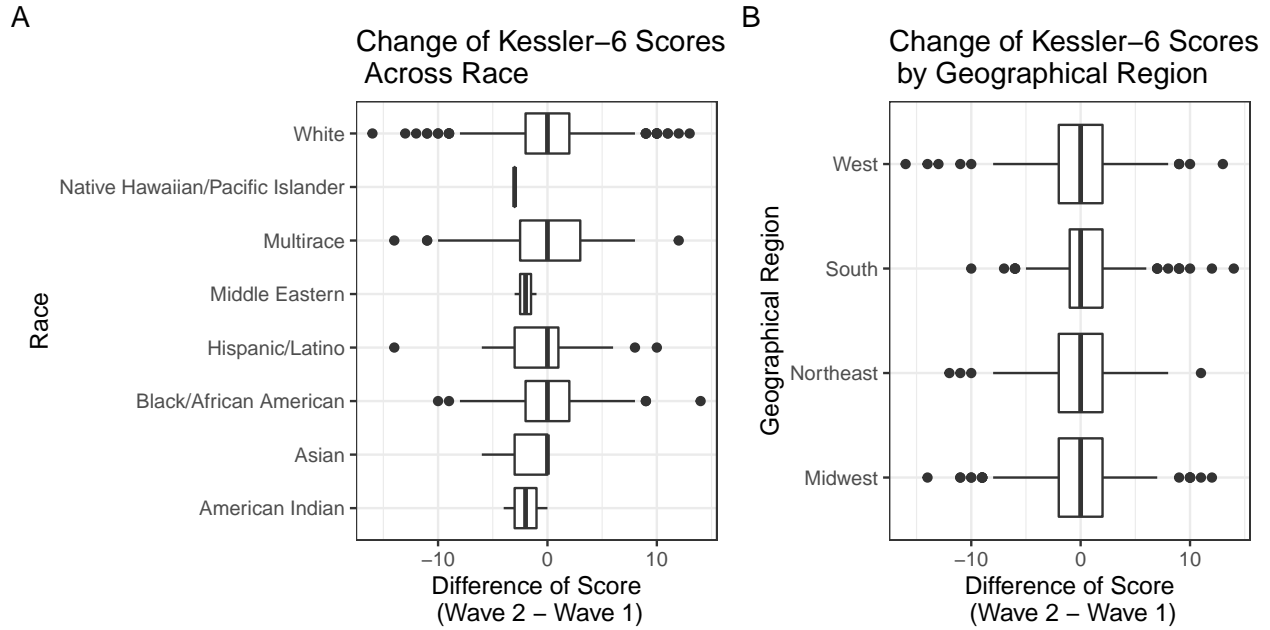**Change of Kessler−6 Scores**
by Geographical Region

Figure 3: Distribution of changes in Kessler−6 scores from Wave 1 survey to Wave 2 survey amongst participants
with varying demographic characeristics of racial identity and geographic region of residence.

From Figure 1 above, we can observe that the distribution of difference in Kessler-6 score is approximately normal amongst the sample of participants that participated in both waves of the Generations survey.

From Figure 2A, we can see that the median change in Kessler-6 Score for individuals who live below the FPL appears to be lower than the median change in Kessler-6 Score for individuals who live above the FPL. Thus, we might want to consider poverty status as a covariate in our model to see whether there is a statistically significant association between poverty status and change in Kessler-6 Score.

From Figure 2B, we can see that the median change in Kessler-6 Score for individuals across all of the age cohorts appears to be relatively the same. However, we would need to include age cohort as a covariate in our model to see whether there is a statistically significant association between age cohort and change in Kessler-6 Score.

From Figure 3A, we can see that the median change in Kessler-6 Score seems to be different for different racial subgroups of the survey sample. From the plot, it appears that Multiracial, White, Hispanic/Latino, Black/African American, and Asian individuals have approximately the same median change in Kessler-6 Score. However, it appears that Native Hawaiian/Pacific Islander and Middle Eastern individuals have a lower median Kessler-6 Score than individuals in all of the other racial categories. Due to the apparent differences in the distributions for change in Kessler-6 Score across the different racial groups, we might want to consider race as a covariate in our model to see whether there is a statistically significant association between race and change in Kessler-6 Score.

From Figure 3B, we can see that the median change in Kessler-6 Score seems to be roughly the same across all of the various geographic census regions. However, while all of the regions seem to have fairly symmetric distributions, the South seems to be skewed left, perhaps implying that more individuals in the South region had a positive change in Kessler-6 score. Given that even though the median is roughly the same, but the distributions vary, it might be useful to include geogrphic region as a covariate in our model to assess whether it has a statistically significant association with change in Kessler-6 Score.

4

Methodology

Model Selection

For modeling the relationship between the covariates of interest and the difference in Kessler-6 outcomes between the Wave 1 survey and Wave 2 survey, I decided to utilize a mixed effects linear regression model. The rationale behind this modeling approach is that while all of the covariates of interest are categorical variables, some have only two factors, such as poverty status, whereas others have many more, such as race. Furthermore, the variables that have many more covariates, such as race, seem to have imbalanced counts of different categories, which is where having random intercepts by race could be useful in modeling the association between racial category and change in Kessler-6 Score.

Model Specification

The model that I ran is specified as follows:

$$Y_i = \beta_0 + \beta_1(\text{Sexuality}_i = \text{Bisexual}) + \beta_2(\text{Sexuality}_i = \text{Other Sexual Minority}) + \beta_3(\text{Education}_i = \text{Some college})$$
$$+ \beta_4(\text{Education}_i = \text{College}) + \beta_5(\text{Education}_i = \text{More than college}) + \beta_6(\text{Urbanicity}_i = \text{Urban})$$
$$+ \beta_7(\text{Region}_i = \text{Midwest}) + \beta_8(\text{Region}_i = \text{South}) + \beta_9(\text{Region}_i = \text{West}) + \beta_{10}(\text{Age Cohort}_i = \text{Middle})$$
$$+ \beta_{11}(\text{Age Cohort}_i = \text{Older}) + \beta_{12}(\text{Sex, Gender}_i = \text{Men, non-transgender})$$
$$+ \beta_{13}(\text{Sex, Gender}_i = \text{Genderqueer/Non-Binary, Female}) + \beta_{14}(\text{Sex, Gender}_i = \text{Genderqueer/Non-Binary, Male})$$
$$+ \beta_{15}(\text{Poverty Status}_i = \text{Below FPL}) + b_{0j} + \epsilon_{ij}$$

$$\text{where } \beta_{0i} \sim N(0, \sigma_0) \perp \epsilon_{ij} \sim N(0, \sigma)$$

with the **brms** default priors set on all parameters as follows:

$$\sigma_0, \sigma, \beta_0 \overset{iid}{\sim} \text{half-Student-t}(3)$$

as well as flat prior specified for all of the β_k coefficients, where $k = 1, \ldots, 15$. The justification for these priors is that they are weakly informative and lead to improved convergence and sampling efficiency for the posterior Gelman (2006). For running the model, I utilized the **brms** package and ran it on `R` version 4.0.3. For the model, I sampled four chains of one-thousand iterations to obtain the posterior samples for the parameters.

Model Diagnostics

I plotted trace plots, density plots, and ACF plots for all of the parameters for all four chains of the model to assess the validity of the model. These can be seen in the **Appendix: Model Validation**.

Results

Table 1: Summary of Parameter Posterior Distributions

Variables	Posterior Mean	2.5th Percentile	97.5th Percentile
Intercept	0.292	-1.632	2.172
Sexual Identity Coefficients			
Bisexual	-0.148	-0.818	0.507
Other Sexual Minority	0.155	-0.778	1.077
Education Coefficients			
Some College	-0.470	-1.259	0.322
College	-0.343	-1.084	0.394
More than College	-0.008	-0.653	0.655
Urbanicity Coefficients			
Urban	0.104	-0.699	0.919
Geographical Region Coefficients			
Midwest	0.021	-0.849	0.867
South	0.925	0.160	1.670
West	0.199	-0.559	0.956
Age Cohort Coefficients			
Middle Age Cohort	0.121	-0.536	0.803
Older Age Cohort	0.088	-0.620	0.776
Sex and Gender Coefficients			
Men, Non-transgender	-1.741	-3.797	0.295
Genderqueer Non-binary, Female	-0.767	-2.126	0.585
Genderqueer Non-binary, Male	-0.560	-1.895	0.773
Poverty Status Coefficients			
Below FPL	-0.406	-1.231	0.405
Random Intercepts by Race			
Sigma for Race Intercept	0.430	0.018	1.360
Asian	-0.031	-1.257	1.076
Black/African American	-0.094	-1.459	0.892
Hispanic/Latino	0.137	-0.556	1.170
Middle Eastern	-0.100	-0.935	0.603
Native Hawaiian/Pacific Islander	-0.060	-1.230	0.923
White	-0.016	-0.739	0.904
American Indian	-0.049	-1.206	1.018
Multiracial	0.249	-0.292	1.158

From the results provided above in Table 1, the only statistically significant association found within the model is that of the association between an LGB individual being located in the South as opposed to the North. In particular, the model results indicate that, holding all other covariates constant, a LGB individual located in the South is expected to have an difference in Kessler-6 Score that is 0.941 units greater than a LGB individual located in the North.

With regards to the differences amongst racial groups, none of the mean differences between the random intercepts by race, conditional on all other covariates being constant, were found to be statistically significant when examining either 66% posterior intervals of 95% posterior intervals, as seen in Table 2 and Table 3 above.

Table 2: Summary of Conditional Means of Comparisons of Race Intercepts (66th Percentile Intervals)

Race Comparison	Lower Quantile	Conditional Mean	Upper Quantile
Hispanic/Latino - Black/African American	-0.6399200	-0.1355634	0.1143662
Middle Eastern - Black/African American	-0.6850692	-0.0590743	0.2635123
Native Hawaiian/Pacific Islander - Black/African American	-0.6595583	-0.0610463	0.2318892
Multiracial - Black/African American	-0.5187746	-0.0889076	0.1826628
Hispanic/Latino - American Indian	-0.4950829	-0.0351645	0.3194161
Asian - American Indian	-0.5143460	-0.0098240	0.3974288
Native Hawaiian/Pacific Islander - Multiracial	-0.4439333	0.0004029	0.3713648
Middle Eastern - American Indian	-0.4935281	0.0007360	0.4498564
Native Hawaiian/Pacific Islander - American Indian	-0.4699744	-0.0003209	0.4437542
Hispanic/Latino - Asian	-0.4533483	-0.0181933	0.3753948
Native Hawaiian/Pacific Islander - Middle Eastern	-0.4395217	-0.0001343	0.4514223
Multiracial - American Indian	-0.4153679	-0.0054221	0.4098787
Middle Eastern - Asian	-0.4108153	0.0087689	0.5064983
Middle Eastern - Hispanic/Latino	-0.3705728	0.0232878	0.4936244
Multiracial - Middle Eastern	-0.4005346	-0.0001392	0.4456568
Native Hawaiian/Pacific Islander - Asian	-0.4348230	0.0046183	0.4696559
Native Hawaiian/Pacific Islander - Hispanic/Latino	-0.3338061	0.0269793	0.4730337
Multiracial - Asian	-0.3586104	0.0060248	0.4671535
Multiracial - Hispanic/Latino	-0.2264872	0.0367057	0.4239932
White - Black/African American	-0.1543984	0.0761175	0.4092623
Black/African American - American Indian	-0.2470383	0.0612530	0.6093463
Black/African American - Asian	-0.2116201	0.0757136	0.6809757
White - Multiracial	-0.0308278	0.2103458	0.6021236
White - American Indian	-0.1123654	0.1630189	0.7098051
White - Native Hawaiian/Pacific Islander	-0.0998664	0.1663578	0.7770739
White - Middle Eastern	-0.1228903	0.1598510	0.7864086
White - Asian	-0.0874132	0.1754477	0.7920921
White - Hispanic/Latino	-0.0041891	0.2702552	0.7392058

Table 3: Summary of Conditional Means of Comparisons of Race Intercepts (95th Percentile Intervals)

Race Comparison	Lower Quantile	Conditional Mean	Upper Quantile
Hispanic/Latino - Black/African American	-1.3091861	-0.1355634	0.5309980
Middle Eastern - Black/African American	-1.9259455	-0.0590743	0.9540543
Native Hawaiian/Pacific Islander - Black/African American	-1.7111213	-0.0610463	0.9096108
Multiracial - Black/African American	-1.1230980	-0.0889076	0.6202730
Hispanic/Latino - American Indian	-1.3820832	-0.0351645	1.2516428
Asian - American Indian	-1.9611206	-0.0098240	1.4772724
Native Hawaiian/Pacific Islander - Multiracial	-1.3998992	0.0004029	1.1694476
Middle Eastern - American Indian	-1.7409567	0.0007360	1.4515960
Native Hawaiian/Pacific Islander - American Indian	-1.5795671	-0.0003209	1.4768364
Hispanic/Latino - Asian	-1.2600602	-0.0181933	1.5729338
Native Hawaiian/Pacific Islander - Middle Eastern	-1.4942181	-0.0001343	1.6424574
Multiracial - American Indian	-1.1440104	-0.0054221	1.4576957
Middle Eastern - Asian	-1.4794773	0.0087689	1.6103841
Middle Eastern - Hispanic/Latino	-1.2649754	0.0232878	1.2986713
Multiracial - Middle Eastern	-1.1628847	-0.0001392	1.6288912
Native Hawaiian/Pacific Islander - Asian	-1.4493303	0.0046183	1.8981972
Native Hawaiian/Pacific Islander - Hispanic/Latino	-1.1744166	0.0269793	1.3012203
Multiracial - Asian	-1.0132617	0.0060248	1.7417123
Multiracial - Hispanic/Latino	-0.6686856	0.0367057	0.9555923
White - Black/African American	-0.5874450	0.0761175	0.8348333
Black/African American - American Indian	-0.9183495	0.0612530	1.7414661
Black/African American - Asian	-0.9129583	0.0757136	2.2005401
White - Multiracial	-0.3051236	0.2103458	1.0317612
White - American Indian	-0.6974109	0.1630189	1.7849452
White - Native Hawaiian/Pacific Islander	-0.6810130	0.1663578	1.7267549
White - Middle Eastern	-0.7441184	0.1598510	1.9821205
White - Asian	-0.6251557	0.1754477	2.2422339
White - Hispanic/Latino	-0.2376619	0.2702552	1.2536358

Discussion

Investigative Conclusions

From the model results, with regards to the general question of whether mental health outcomes differ across LGB individuals from various different demographic background, it appears that the only significant association with difference in outcome is between those LGB individuals located in the South as opposed to the North. This finding struck me as somewhat surprising given that typically Northern states tend to have better legal protections in place for LGB people and more positive societal attitudes towards LGB people than Southern states (("Laws &Amp; Legislation - Hrc," n.d.)).

With regards to the secondary question of interest of examining whether there are any differences amongst the racial groups, it appears that there are no statistically significant differences in the magnitude of the association between race and mental health outcomes across all of the racial groups. This finding also struck me as surprising, as it seems counter-intuitive to what has been suggested in the literature elsewhere that other demographic characteristics, such as race, might have differential associations with mental health outcomes within the population LGB people in the U.S. (Bostwick et al. (2014)).

Limitations and Future Work

One major limitation of this study is that the modeling approach used in this paper does not account for survey weights associated with each individual that was sampled in the survey. Within the survey, each respondent was assigned a survey weight in order to correct bias that would ensue from the sample due to the fact that the sample was not a representative sample of the LGB population at large in the U.S. In particular, this sample oversampled White LGB individuals and had very few observations for LGB people from other racial groups such as Native Hawaiian/Pacific Islander, Asian, and Middle Eastern. Thus, it is possible that not accounting for the survey weights in the modeling approach has led to biased results in the model findings. In future work, researchers could try and incorporate survey weights into the modelling approach, as perhaps that would remove sources of bias when sampling from the posterior distributions of the model parameters. Such approaches might reference work done in Si et al. (2017) to inform the modeling approach.

Another way to approach the modeling that I had considered to alleviate the issue of survey weights was utilizing a frequentist linear mixed effects model that could account for survey weights. However, it is difficult to quantify uncertainty for parameter estimates using frequentist linear mixed effects models, which could lead to results that are not readily interpretable in the context of the research objectives.

Another limitation of this paper is that I utilized a linear model that can technically have non-integer outcomes for the difference in Kessler-6 score even though the Kessler-6 scale is technically only an integer-value scale and thus can only have integer-value outcomes for the difference in score. Thus, another idea for future work could be to instead model the outcome as a binary categorical response of whether or not the change in score was positive or negative. However, this would lead to loss of information on the magnitude of the change that is associated with the covariates.

Summary

Overall, the findings of this paper are that the only covariate that seems to be have a statistically significant association with change in Kessler-6 Score is that of whether a LGB individual is from the South as opposed to being from the Northeast. In particular, this association is in the positive direction, which seems to indicate that perhaps LGB individuals in the South had improved mental health outcomes over the course of the Generations survey as opposed to LGB individuals from the Northeast. The other question of major interest was whether there were any differential impacts of race on mental health outcomes of LGB individuals, and from the modeling results in this paper, there do not appear to be any statistically significant differences in how race impacts mental health outcomes across the various racial groups that were present in the survey sample. Thus, we cannot make any conclusive evidence of an association with belonging to different racial groups having an impact on mental health outcomes.

Appendix

Model Validation

Trace Plots

In order to assess the validity of the model, I plotted trace plots for all four chains that were run for the model. As shown below, the trace plots indeed appear to converge.
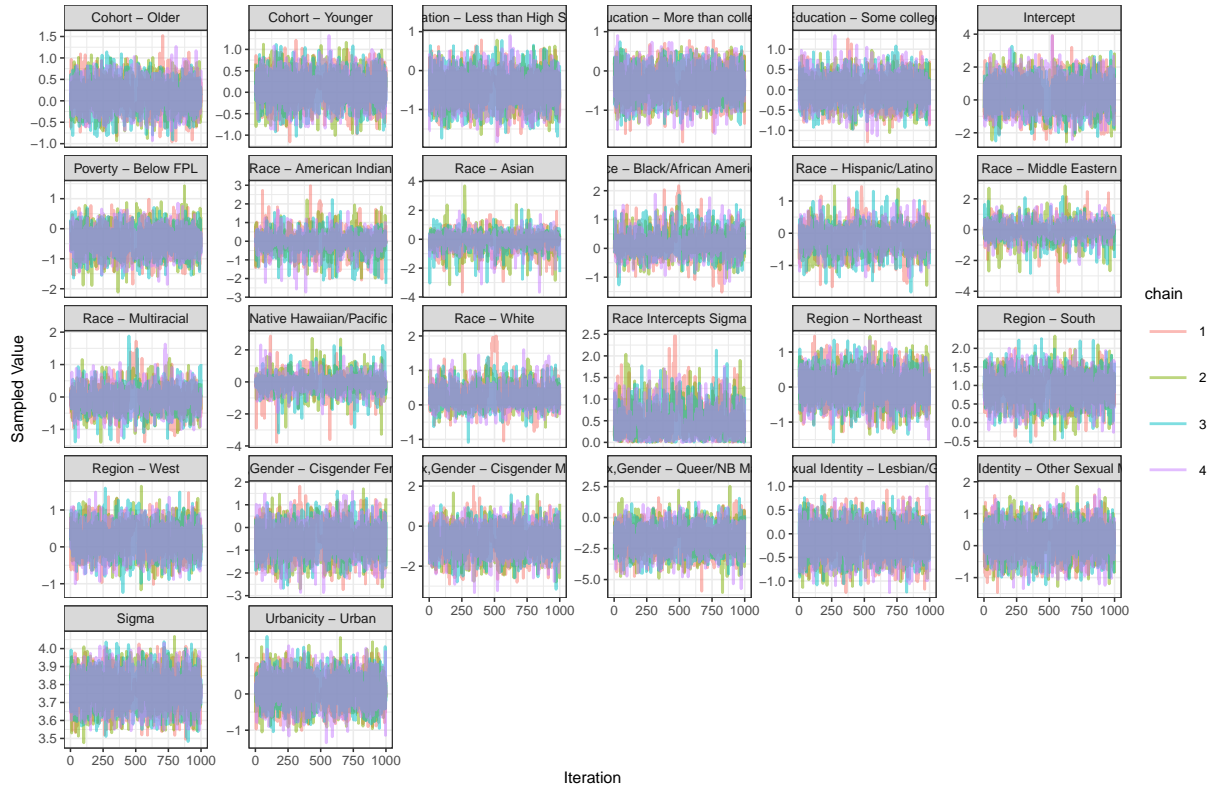

Figure A–1: Trace Plots for all parameters.

Density Plots

In order to assess the validity of the model, I plotted density plots for the potsterior distributions of all the parameters for all four chains that were run for the model. As shown below, the density plots for the coefficient terms appear to be approximately normal, and the trace plots for the standard deviation of the race random intercepts as well as the standard deviation of the residuals both appear to be approximately half-Student-t distributions, as expected. Thus, the posterior densitites seem to appear as we would expect them to appear.
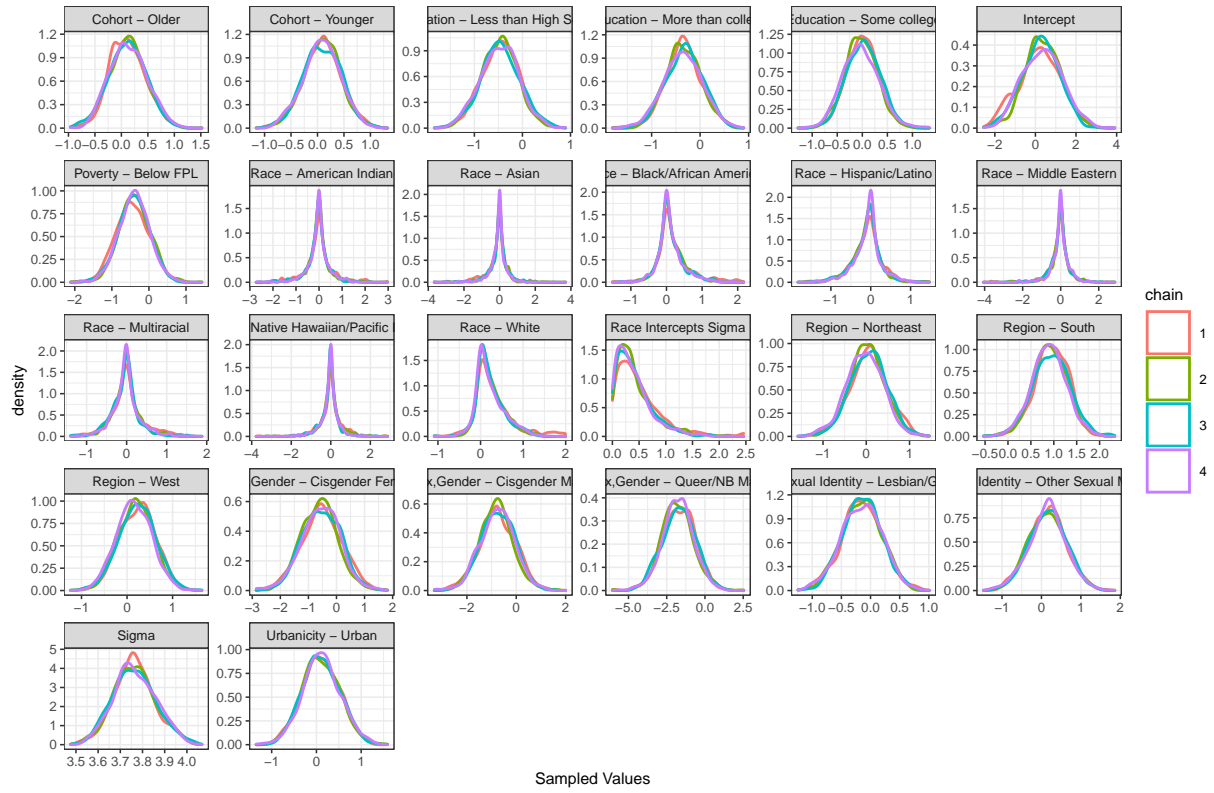
Figure A–2: Density Plots for all parameters.

ACF Plots

I plotted ACF plots to check the correlation between samples taken from subsequent iterations of the model to assess whether the posterior samples sufficiently explored the posterior distribution space. As Figures A-3 through A-7 below show, the ACF is fairly low for every 20 iterations, and thus it appears that posterior space was sufficiently explored.
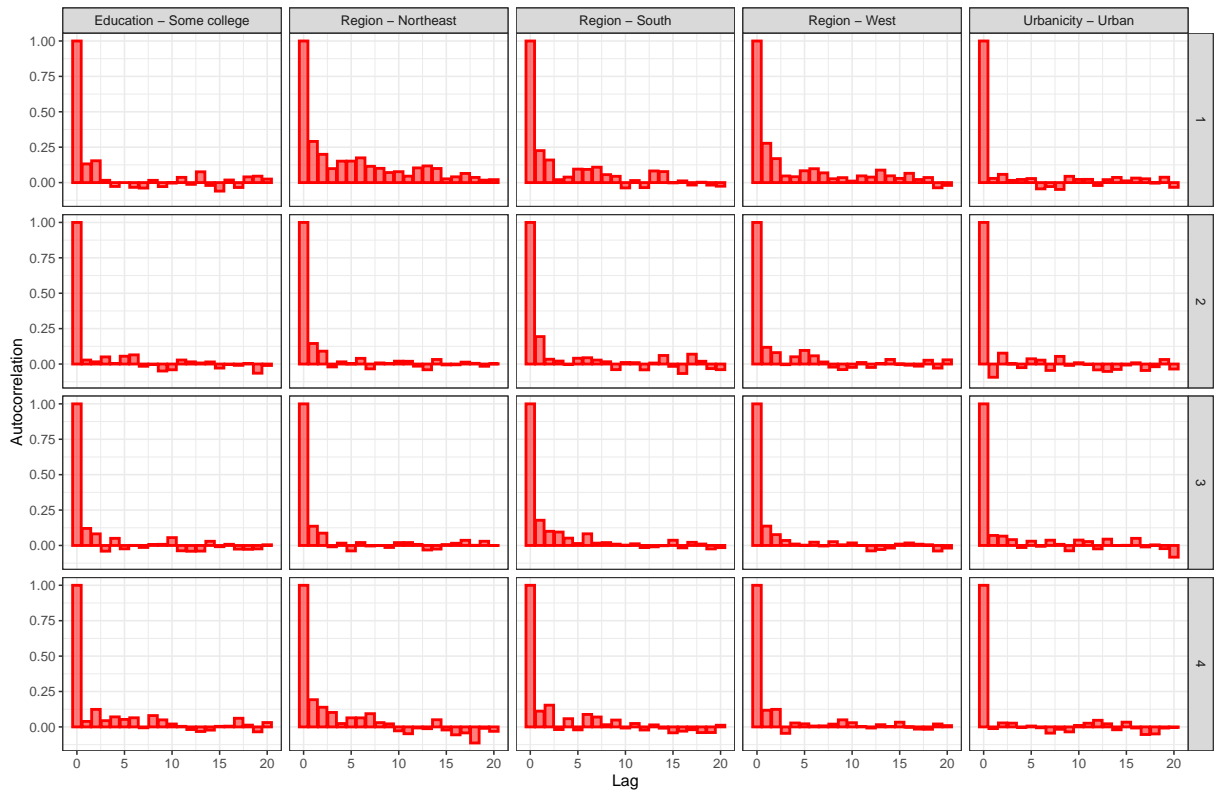
Figure A–3: ACF Plots for parameters.
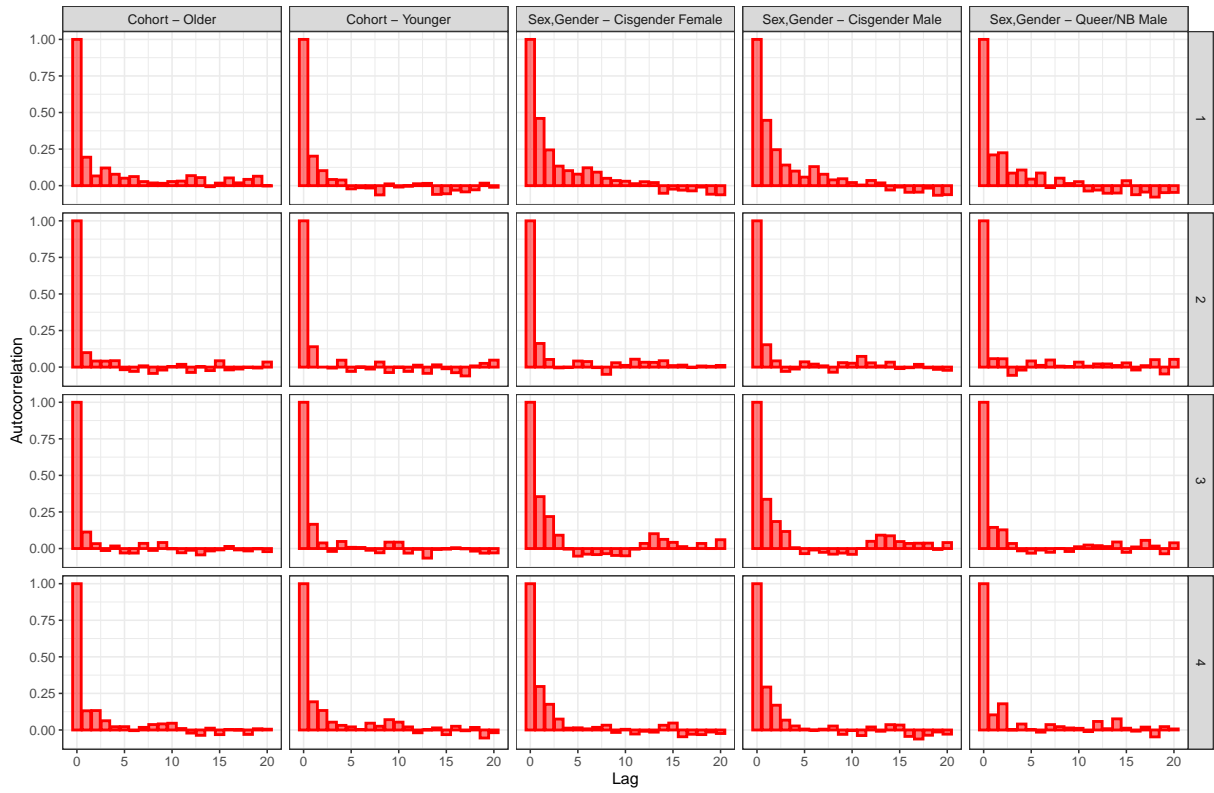

Figure A–4: ACF Plots for parameters.

Figure A–5: ACF Plots for parameters.
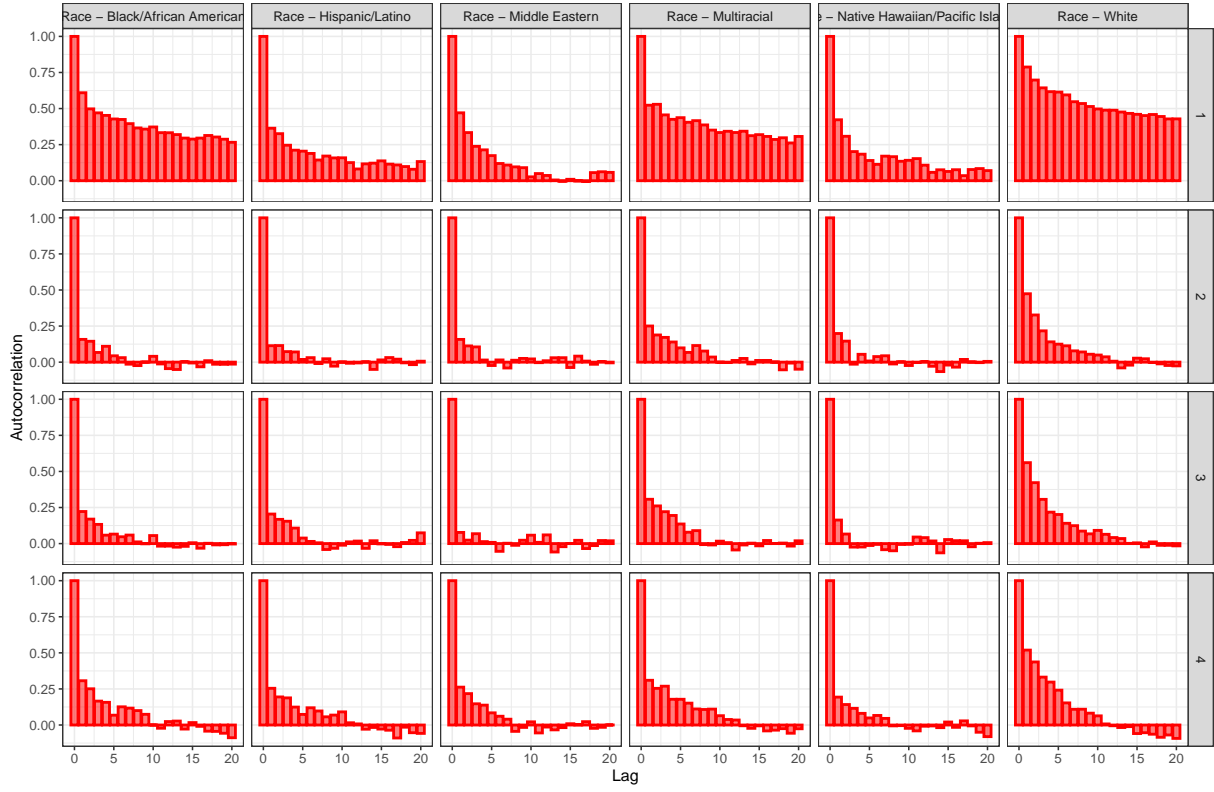

Figure A–6: ACF Plots for parameters.

Figure A–7: ACF Plots for parameters.

Plot of Conditional Mean of Posterior Comparisons for Race Intercepts

The plot below shows visually the results presented earlier in Tables 2 and 3. As seen earlier, none of the intervals seem to not contain zero, thus leading to us being unable to conclude if there are any differential impacts of race on mental health outcomes amongst LGB individuals.
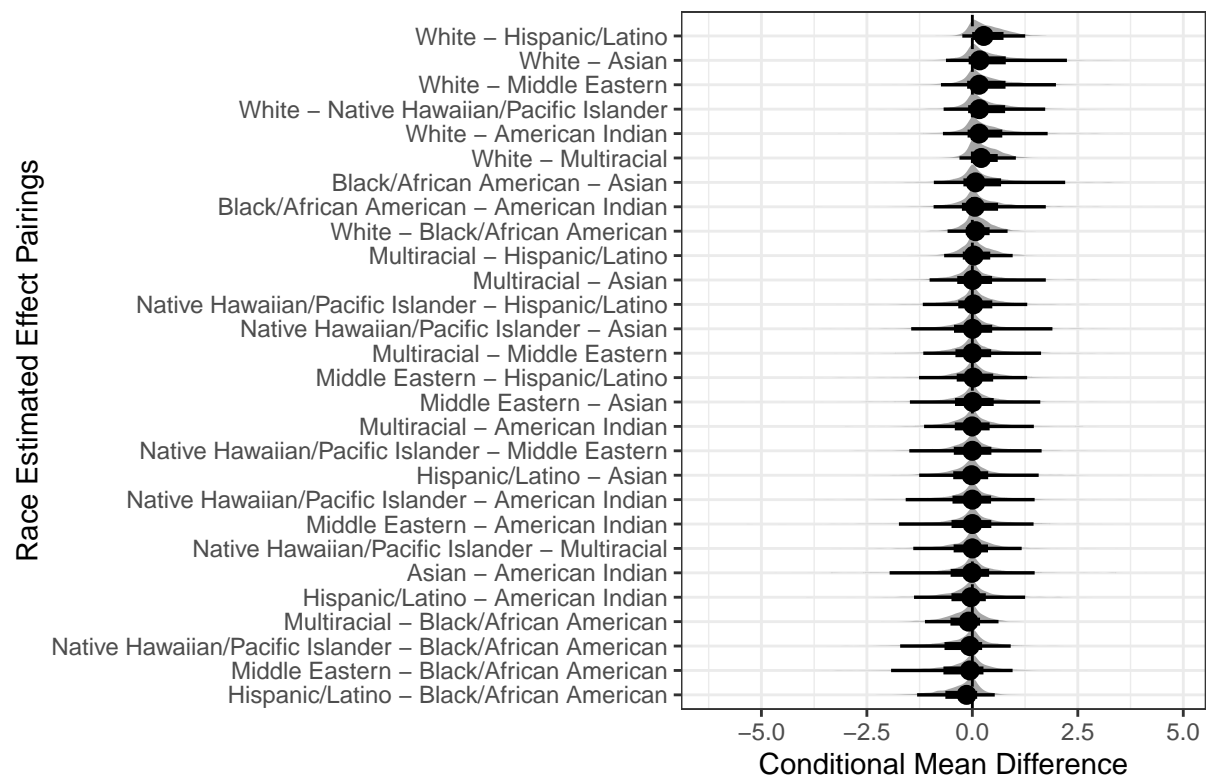
Figure A–8: Plot of conditional mean differences across random intercepts by race.

References

Bostwick, Wendy B., Carol J. Boyd, Tonda L. Hughes, Brady T. West, and Sean Esteban McCabe. 2014. "Discrimination and Mental Health Among Lesbian, Gay, and Bisexual Adults in the United States." *American Journal of Orthopsychiatry* 84 (1): 35–45. https://search.ebscohost.com/login.aspx?direct=true&db=pdh& AN=2014-06035-006&site=ehost-live&scope=site.

Gelman, Andrew. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models." *Bayesian Analysis* 1 (3). https://doi.org/10.1214/06-ba117a.

Kessler, Ronald C., Patricia Berglund, Olga Demler, Robert Jin, Doreen Koretz, Kathleen R. Merikangas, A. John Rush, Ellen E. Walters, and Philip S. Wang. 2003. "The Epidemiology of Major Depressive DisorderResults From the National Comorbidity Survey Replication (NCS-R)." *JAMA* 289 (23): 3095–3105. https://doi.org/10.1001/jama.289.23.3095.

"Laws &Amp; Legislation - Hrc." n.d. *Human Rights Campaign.* https://www.hrc.org/resources/laws-legislation.

Meyer, Ilan H. 2003. "Prejudice, Social Stress, and Mental Health in Lesbian, Gay, and Bisexual Populations: Conceptual Issues and Research Evidence." *Psychological Bulletin* 129 (5): 674–97. https://www.ncbi.nlm. nih.gov/pmc/articles/PMC2072932/?report=classic.

Si, Yajuan, Rob Trangucci, Jonah Sol Gabry, and Andrew Gelman. 2017. "Bayesian Hierarchical Weighting Adjustment and Survey Inference." arXiv. https://doi.org/10.48550/ARXIV.1707.08220.