Final

Reyna Orellana

5/12/2021

#Introduction I am interested in abortion opinions within the United States. I want to explore how notions of whether women should be legally allowed abortions, moral opposition, and reasons for abortion vary by social status categories measured by factors like race, gender, religion, class, age, income, and education. Research question that guided this paper are the following: What are the general opinions regarding abortion? Does class impact abortion perceptions? What other factors impact abortion perceptions? Reproduction rights and access to health clinics have been of major study. I am interested in understanding the most current data on this topic. It might shed light in areas that seems intuitive but also it might show that there needs to be a better understanding of changes over time. Although in hindsight I should have downloaded different years to do this type of comparison. I would have wanted to download maybe 5 or 10 year increments to do a time analysis of particular topics within abortion and see if they changed. However I only downloaded the latest year that the General Social Survey has released which is from 2018.

Adding packages and data summary

#Data and Methods

For this analysis we will be using a General Social Survey (GSS) dataset from 2018. Using GSS allows us to generalize findings to the broader United States population because of its random sampling methods. The GSS contains a standard core of demographic, behavioral, and attitudinal questions, plus other topics of special interest.

Of interest to us, I kept 36 variables to analyze, 10 constructed variables, and will present different methods for the use of this dataset. The abortion variables are the following: abany (abortion if woman wants for any reason 1 yes 2 no), abdefect (strong chance of serious defect 1 yes 2 no), abfelegal (women only: women should be able to have legal abortions 1 yes 2 no), abhelp1 (r would help with arrangements for abortion 1 yes 2 no), abhelp2 (r would help with paying for abortion 1 yes 2 no), abhelp3 (r would help with paying for abortion-related other costs 1 yes 2 no), abhelp4 (r would help with emotional support for abortion 1 yes 2 no), abhlth (woman's health seriously endangered 1 yes 2 no), abinspay (should people be able to use health insurance for abortion 1 yes 2 no), abmedgov1 (abortion: woman and doctor or govt should decide what info needed 1 yes 2 no), abmedgov2 (Childbirth: woman and doctor or govt should decide what info is needed 1 yes 2 no), abmlegl, (men only: women should be able to have legal abortions 1 yes 2 no), abmoral (r has a moral opposition to abortion 1 yes 2 no), abnomore (married—wants no more children 1 yes 2 no), abpoor (low income—can't afford more children 1 yes 2 no), abpoorw (wrong for woman to get abortion if low income? 1 always wrong 2 almost always wrong 3 wrong only sometimes 4 not wrong at all), abrape (pregnant as result of rape 1 yes 2 no), absingle (not married 1 yes 2 no), abstate1 (difficulty of obtaining abortion in r's state 1very easy 2 - easy 3 - neither easy or hard 4 - hard, 5- very hard. The GSS codebook provides the extended version of the questions that respondents were asked. For our methods, we are using regression analysis because we want to examine the relationship between our variables of interest abortion attitudes and racial differences. I constructed one scale to determine abortion attitudes. First, I recoded all 4 variables to make sure high values indicate the more positive attitudes toward abortion. I also used chi square tests to determine differences in the conditional distribution of abortion on another variable. There was no significance difference for being morally opposed to abortion by racial categories or gender. So I reject my alternative hypothesis because moral opposition does not depend on race or gender. Although this might have been because abmoral is not the right type of dv since its categories are only 3 categories that cannot be

added up and divided. However in the following model I used two different variables that were better suited. After creating a variable that varied from 10 being strong yes help to 1 being strong no help for abortion and class. My hypothesis is that there will be no statistical significance. My alternative hypothesis is that there will be statistical significance. In this model there was a significance difference in the help that participants would give to a family member or friend that decided to have an abortion and this changes by class. In this model I reject my null hypothesis because there is statistical significance.

```
#recoding variables from numeric to categorical.
#race
ab18<- ab18 %>%
  mutate(race_3cat = case_when(
    race == 1 ~ "White",
    race == 2 ~ "Black",
    race == 3 ~ "Other"
  ))
#checking
table(select(ab18, race, race_3cat ))
```

```
##     race_3cat
## race Black Other White
##    1     0     0  1120
##    2   252     0     0
##    3     0   187     0
```

```
#age
summary(ab18$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   18.00   35.00   50.00   50.12   64.00   89.00       7
```

```
#recoding age variable
ab18 <- ab18%>%
  mutate(age_4cat = case_when( age >= 18 & age <= 29 ~ "18-29",
                               age >= 30 & age <= 39 ~ "30-39",
                               age >= 40 & age <= 49 ~ "40-49",
                               age >= 50 & age <= 89 ~ "50+"))
table(select(ab18, age_4cat))
```

```
##
## 18-29 30-39 40-49   50+
##   235   275   255   787
```

```
#checking if recoded correctly
table(select(ab18, age_4cat, age))
```

```
##         age
## age_4cat 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
##    18-29 13 16  6 16 18 14 23 24 20 28 28 29  0  0  0  0  0  0  0  0  0  0  0
##    30-39  0  0  0  0  0  0  0  0  0  0  0  0 33 20 29 27 40 25 21 25 22 33  0
##    40-49  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 28
##    50+    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##         age
## age_4cat 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63
##    18-29  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    30-39  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    40-49 27 27 26 24 23 30 23 24 23  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    50+    0  0  0  0  0  0  0  0  0 24 27 22 38 24 32 33 30 28 27 28 28 16 27
```

```
##          age
## age_4cat 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86
##    18-29  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    30-39  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    40-49  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##    50+   25 27 28 17 22 20 33 27 15 18 18 17 13 22  7  8 11  8 10  6  5  8 10
##          age
## age_4cat 87 88 89
##    18-29  0  0  0
##    30-39  0  0  0
##    40-49  0  0  0
##    50+    2  5 21
```

```
#able to have an abortion legally or not. Merging abmlegal and abflegal
ab18 <- ab18 %>%
  mutate(legal = case_when(abfelegl == 1 | abmelegl == 1 ~ 1,
                           abfelegl == 2 | abmelegl == 2 ~ 2,
                           abfelegl == 3 | abmelegl == 3 ~ 3))
ab18 <- ab18 %>%
  mutate(legal_cat = case_when(abfelegl == 1 | abmelegl == 1 ~ "Should",
                               abfelegl == 2 | abmelegl == 2 ~ "Should not",
                               abfelegl == 3 | abmelegl == 3 ~ "Depends"))
#checking
table(select(ab18, legal))
```

```
##
##   1   2   3
## 601 222 689
```

```
#class 1- lower class 2- working class 3- middle class 4- upper class
table(select(ab18, class))
```

```
##
##   1   2   3   4
## 149 654 695  53
```

```
ab18 <- ab18 %>%
  mutate(classr = case_when(class == 1 ~"lower class",
                            class == 2 ~ "working class",
                            class == 3 ~ "middle class",
                            class == 4 ~ "upper class"))
table(select(ab18, classr))
```

```
##
##   lower class  middle class   upper class working class
##           149           695            53           654
```

```
#recoding help abotion questions. Will add 4 of the help variables abhelp1-abhelp4
ab18<- ab18 %>%
  mutate(abortion_help = case_when(
    abhelp1 == 1 & abhelp2 == 1 & abhelp3 == 1 & abhelp4 ==1 ~ "Strong Yes",
    abhelp1 == 1 & abhelp2 == 2 & abhelp3 == 1 & abhelp4 ==1 ~ "Weak Yes",
    abhelp1 == 1 & abhelp2 == 1 & abhelp3 == 2 & abhelp4 ==1 ~ "Weak Yes",
    abhelp1 == 1 & abhelp2 == 1 & abhelp3 == 1 & abhelp4 ==2 ~ "Weak Yes",
    abhelp1 == 2 & abhelp2 == 1 & abhelp3 == 1 & abhelp4 ==1 ~ "Weak Yes",
    abhelp1 == 1 & abhelp2 == 2 & abhelp3 == 2 & abhelp4 ==2 ~ "Weak No",
    abhelp1 == 2 & abhelp2 == 2 & abhelp3 == 2 & abhelp4 ==1 ~ "Weak No",
```

```
    abhelp1 == 2 & abhelp2 == 1 & abhelp3 == 2 & abhelp4 ==2 ~ "Weak No",
    abhelp1 == 2 & abhelp2 == 2 & abhelp3 == 1 & abhelp4 ==2 ~ "Weak No",
    abhelp1 == 2 & abhelp2 == 2 & abhelp3 == 2 & abhelp4 ==2 ~ "Strong No",
    TRUE ~ "Not Available"
  ))
#different way to code abortion help
table(select(ab18, abortion_help))
```

```
##
## Not Available     Strong No    Strong Yes       Weak No      Weak Yes
##           375           150           420           227           387
```

```
#recoding help abotion questions. Will add 4 of the help variables abhelp1-abhelp4 on a scale of 10
ab18<- ab18 %>%
  mutate(abhelpr = case_when(abhelp1 == 1 & abhelp2 == 1 & abhelp3 == 1 & abhelp4 ==1 ~ 10,
                             abhelp1 == 1 & abhelp2 == 2 & abhelp3 == 1 & abhelp4 ==1 ~ 9,
                             abhelp1 == 1 & abhelp2 == 1 & abhelp3 == 2 & abhelp4 ==1 ~ 8,
                             abhelp1 == 1 & abhelp2 == 1 & abhelp3 == 1 & abhelp4 ==2 ~ 7,
                             abhelp1 == 2 & abhelp2 == 1 & abhelp3 == 1 & abhelp4 ==1 ~ 6,
                             abhelp1 == 1 & abhelp2 == 2 & abhelp3 == 2 & abhelp4 ==2 ~ 5,
                             abhelp1 == 2 & abhelp2 == 2 & abhelp3 == 2 & abhelp4 ==1 ~ 4,
                             abhelp1 == 2 & abhelp2 == 1 & abhelp3 == 2 & abhelp4 ==2 ~ 3,
                             abhelp1 == 2 & abhelp2 == 2 & abhelp3 == 1 & abhelp4 ==2 ~ 2,
                             abhelp1 == 2 & abhelp2 == 2 & abhelp3 == 2 & abhelp4 ==2 ~ 1))
#different way to code abortion help
table(select(ab18, abhelpr))
```

```
##
##   1   2   4   5   6   7   8   9  10
## 150   1 208  18   3   1  18 365 420
```

```
ab18<- ab18 %>%
  mutate(abhelp = case_when(
    abhelp1 == 1 & abhelp2 == 1 & abhelp3 == 1 & abhelp4 ==1 ~ "Yes",
    abhelp1 == 2 & abhelp2 == 2 & abhelp3 == 2 & abhelp4 ==2 ~ "No"
  ))
table(select(ab18, abhelp))
```

```
##
##  No Yes
## 150 420
```

```
#sex (56% 868 female, 44% 691 male)
ab18<- ab18 %>%
  mutate(sexp = case_when(sex == 1 ~ "Male",
                          sex == 2 ~ "Female"
  ))

table(select(ab18, sexp))
```

```
##
## Female   Male
##    868    691
```

```
prop.table(table(select(ab18, sexp)))
```

```
##
```

```
##    Female      Male
## 0.5567672 0.4432328
```

```
table(select(ab18, income))
```

```
##
##   1   2   3   4   5   6   7   8   9  10  11  12
##  22  21  12  13   7   6   8  28  95  50  98 960
```

```
prop.table(table(select(ab18, income)))
```

```
##
##           1           2           3           4           5           6
## 0.016666667 0.015909091 0.009090909 0.009848485 0.005303030 0.004545455
##           7           8           9          10          11          12
## 0.006060606 0.021212121 0.071969697 0.037878788 0.074242424 0.727272727
```

```
#I want to understand my dataset more so I will do some tables
#abany abortion for any reason (1 is Yes 50% and 2 is No 50%)
table(select (ab18, abany))
```

```
##
##   1   2
## 381 377
```

```
prop.table(table (select(ab18, abany)))
```

```
##
##         1         2
## 0.5026385 0.4973615
```

```
#abdefect (1 is Yes 77% and 2 is No 23%) only women were asked this Q
prop.table(table(select(ab18, abdefect)))
```

```
##
##         1         2
## 0.7743658 0.2256342
```

```
#abfelegl (1 is should 38% 2 is should not 15% 3 is depends 46%)
table(select(ab18, abfelegl))
```

```
##
##   1   2   3
## 329 126 387
```

```
prop.table(table(select(ab18, abfelegl)))
```

```
##
##         1         2         3
## 0.3907363 0.1496437 0.4596200
```

```
#abinspay
table(select(ab18, abinspay))
```

```
##
##   1   2
## 718 706
```

```
#abmelegl (1 is should 41% 2 is should not 13% 3 is depends 45%)
table(select(ab18, abmelegl))
```

```
##
##   1   2   3
## 272  96 302
```

```
prop.table(table(select(ab18, abmelegl)))
```

```
##
##         1         2         3
## 0.4059701 0.1432836 0.4507463
```

```
#abmoral 1- morally opposed (28%) 2-not morally opposed (27%) 3- it depends (45%)
table(select(ab18, abmoral))
```

```
##
##   1   2   3
## 425 415 681
```

```
prop.table(table(select(ab18, abmoral)))
```

```
##
##         1         2         3
## 0.2794214 0.2728468 0.4477318
```

```
#abhlth 1 - yes (91%) 2 - no (9%)
prop.table(table(select(ab18, abhlth)))
```

```
##
##          1          2
## 0.90848806 0.09151194
```

```
#abmedgov1 1 - yes (91%) 2 - no (9%)
prop.table(table(select(ab18, abmedgov1)))
```

```
##
##          1          2
## 0.92980132 0.07019868
```

```
#abmedgov2 1 - yes (93%) 2 - no (7%)
prop.table(table(select(ab18, abmedgov2)))
```

```
##
##          1          2
## 0.94722598 0.05277402
```

```
#abnomore 1 - yes (95%) 2 - no (5%)
prop.table(table(select(ab18, abnomore)))
```

```
##
##         1         2
## 0.5193591 0.4806409
```

```
#abpoor 1 - yes (52%) 2 - no (48%)
prop.table(table(select(ab18, abpoor)))
```

```
##
##         1         2
## 0.4966711 0.5033289
```

```
#abpoorw 1 - yes (50%) 2 - no (50%)
prop.table(table(select(ab18, abpoorw)))
```

```
##
##          1          2          3          4
## 0.46036036 0.08378378 0.12072072 0.33513514
```

```
#abrape 1 - always wrong (46%) 2 - almost always wrong (8%) 3 wrong only sometimes (12%) 4 not wrong at
prop.table(table(select(ab18, abrape)))
```

```
##
##         1         2
## 0.7981283 0.2018717
```

```
#absingle 1 - yes (80%) 2 - no (20%)
prop.table(table(select(ab18, absingle)))
```

```
##
##         1         2
## 0.4660453 0.5339547
```

```
#abstate1 1 - very easy (20%) 2 - easy (36%) 3 - neither easy or hard (28%) 4 - hard (11%), 5 - very ha
prop.table(table(select(ab18, abstate1)))
```

```
##
##          1          2          3          4          5
## 0.19480519 0.36444805 0.28165584 0.11282468 0.04626623
```

```
#education
summarise(ab18, coldeg1)
```

```
## # A tibble: 1,559 x 1
##                coldeg1
##              <dbl+lbl>
##  1     1 [associate's]
##  2     2 [bachelor's]
##  3 NA(i) [IAP]
##  4     2 [bachelor's]
##  5 NA(i) [IAP]
##  6 NA(i) [IAP]
##  7     4 [mba]
##  8 NA(i) [IAP]
##  9 NA(i) [IAP]
## 10 NA(i) [IAP]
## # ... with 1,549 more rows
```

```
table(select(ab18, coldeg1))
```

```
##
##   1   2   3   4   5   6   7   8
##  82 215  95   9   4  15   5   7
```

```
#religion 1- protestant 2- catholic 3- jewish 4- none 5- Other 6- buddhism 7- hinduism 8- other eastern
table(select(ab18, relig))
```

```
##
##   1   2   3   4   5   6   7   8   9  10  11  12  13
## 741 360  26 339  20  11   6   1  10   3  26   1   1
```

```
ab18<- ab18%>%
  mutate(religion = case_when(relig == 1 ~ "Protestant",
                              relig == 2 ~ "Catholic",
```

```
                              relig == 3 ~ "Jewish",
                              relig == 4 ~ "None",
                              relig == 5 ~ "Other",
                              relig == 6 ~ "Buddhism",
                              relig == 7 ~ "Hinduism",
                              relig == 8 ~ "Other Eastern religion",
                              relig == 9 ~ "Muslim/Islam",
                              relig == 10 ~ "Orthodox-christian",
                              relig == 11 ~ "Christian",
                              relig == 12 ~ "Native American",
                              relig == 13 ~ "Inter-nondenomentational"
  ))
prop.table(table(select(ab18, religion)))
```

```
##
##                 Buddhism                 Catholic                Christian
##             0.0071197411             0.2330097087             0.0168284790
##                 Hinduism Inter-nondenomentational                    Jewish
##             0.0038834951             0.0006472492             0.0168284790
##             Muslim/Islam          Native American                     None
##             0.0064724919             0.0006472492             0.2194174757
##       Orthodox-christian                    Other   Other Eastern religion
##             0.0019417476             0.0129449838             0.0006472492
##               Protestant
##             0.4796116505
```

```
#chi square test
#inspect cross-tab with column proportions
prop.table(xtabs(~ abmoral + race_3cat, ab18))
```

```
##        race_3cat
## abmoral      Black      Other      White
##       1 0.03418803 0.03024326 0.21499014
##       2 0.04536489 0.03353057 0.19395135
##       3 0.08218277 0.05588429 0.30966469
```

```
#calculate chisq.test
chisq.test(ab18$abmoral,ab18$race_3cat)
```

```
##
##  Pearson's Chi-squared test
##
## data:  ab18$abmoral and ab18$race_3cat
## X-squared = 9.0347, df = 4, p-value = 0.06024
```

```
#There was no significance differnce for being morally opposed to abortion by racial or gender categori
prop.table(xtabs(~ abmoral + sexp, ab18))
```

```
##        sexp
## abmoral    Female      Male
##       1 0.1663379 0.1130835
##       2 0.1505588 0.1222880
##       3 0.2399737 0.2077581
```

```
#calculate chisq.test
chisq.test(ab18$abmoral,ab18$sexp)
```

```
##
##  Pearson's Chi-squared test
##
## data:  ab18$abmoral and ab18$sexp
## X-squared = 3.7906, df = 2, p-value = 0.1503
```

```
##calculate group means of DV for categorical IV
tapply(ab18$abhelpr,ab18$classr,mean, na.rm=TRUE)
```

```
##   lower class  middle class   upper class working class
##      6.747826      7.707182      8.024390      7.095634
```

```
anova(lm(ab18$abhelpr~ab18$classr))
```

```
## Analysis of Variance Table
##
## Response: ab18$abhelpr
##               Df Sum Sq Mean Sq F value   Pr(>F)
## ab18$classr    3    160  53.327  5.1191 0.001604 **
## Residuals   1176  12251  10.417
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#estimate bivariate regression lm(DV~catIV)
summary(lm(ab18$abhelpr~ab18$classr))
```

```
##
## Call:
## lm(formula = ab18$abhelpr ~ ab18$classr)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.024 -3.096  1.904  2.293  3.252
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  6.7478     0.3010  22.420  < 2e-16 ***
## ab18$classrmiddle class      0.9594     0.3313   2.896  0.00385 **
## ab18$classrupper class       1.2766     0.5871   2.174  0.02987 *
## ab18$classrworking class     0.3478     0.3350   1.038  0.29941
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.228 on 1176 degrees of freedom
##   (379 observations deleted due to missingness)
## Multiple R-squared:  0.01289,    Adjusted R-squared:  0.01037
## F-statistic: 5.119 on 3 and 1176 DF,  p-value: 0.001604
```

#Reflection This was a long process to make sure there was a structure that was easy to follow. It started as an exploratory quest but as I narrowed what I thought would be my focus I soon realized that I had assumptions that I could test. I tried first to understand my dataset and see what variables had scales and which one were numeric variables that represent categorical data. I also had to think through what demographic data or social status variables I could use to think about independent and dependent variables. I am concerned with this part of my data because I could think through more about this but I think it requires more time and understanding around how to correctly use certain variables together and when not to. I tried to recode variables that I needed to turn from numeric to categorical. I later found out that most variables were coded the same with 1 being yes and 2 being no. This allowed me to inspect the tables faster.