Relationship between social capital and election results

Anisha Babu[1], Hyeonjin Cha[1], Diana DeWald[1], & Murat Kezer[1]

[1] University of Oregon

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line. Enter author note here.

The authors made the following contributions. Anisha Babu: Conceptualization, Data Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Hyeonjin Cha: Conceptualization, Data Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Diana DeWald: Conceptualization, Data Analysis, Writing - Original Draft Preparation, Writing - Review & Editing; Murat Kezer: Conceptualization, Data Analysis, Writing - Original Draft Preparation, Writing - Review & Editing.

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words "**here we show**" or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: keywords

Word count: X

Relationship between social capital and election results

Introduction

Social science literature has extensively examined the relationship between social capital and politics (e.g. Morales & Guigni, 2016; Jottier & Heyndels, 2012; La Due Lake & Huckfeldt, 1998). However, relatively little is known on the impact of social capital election results.

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Data

- County Presidential Election Returns 2000-2016 (MIT Election Data and Science Lab, 2018)
  - County level returns for presidential elections from 2000 to 2016
  - Election results across the years in one dataset & in tidy format
- The production of social capital in US counties (Rupasingha, Goetz, & Freshwater, 2006, with updates)
  - County level count of various establishments defined by NAICS code
  - Different variables measured across years; new dataset for each year

Data Preparation

Load data and clean names. We first load the datasets and clean the variable names.

Clean data.

Election data.

- We start with the election data as it is more comprehensive in terms of the number of counties. First, we select the variables of interests. Then, we select the election year (i.e., 2000, 2008, 2012, 2016) that we will match with social capital data.

- The name of the year variable is changed in a way that shows it is the year of election (so that it is not mixed with the same year variable in social capital data).

- We create new datasets for each presidential election we are interested in. These will be later merged with corresponding social capital data.

Social capital data.

- For each social capital dataset (i.e., 1997, 2005, 2009, 2014), we first add state code for some counties that do not readily contain that information. Then, we create two variables out of the area name such that we have different variables for county names and state codes.

- We select the relevant variables and clean the variable names.

- We create a year variable indicating when the data were collected.

- Finally, we reorder variables so that the order of the variables is the same across datasets. This will be useful when we want to merge social capital data across year so that we can get descriptive statistics for each year simultaneously and that we can visualize the changes across years in social capital.

Merge Datasets.

- First, we merge social capital data across years for reasons explained above, and call it `s_capital`.
- Next, we merge corresponding election and social capital data for 4 time points. **In doing so,** we keep the rows that exist in both election and social capital data. For instance, if we do not have the election information for a county, we do not include it in the merged dataset even if we have that county's social capital data. These datasets are called `df_year`. *Year* denotes the year of election. Also, we remove the duplicate variables (i.e., state and county names) and fix the names. We did not remove them earlier because we first wanted to merge the social capital data with all the variables.
- Finally, we merge all election and social capital data in the same dataset (i.e., `df`). In addition, we created another dataset using `pivot_wider()` to have variables for the candidate votes per political party. Then, we removed the intermediate objects (i.e., all data frames except for `df` and `df_wide`).

Data analysis

We used R [Version 3.6.1; 8] and the R-packages *dplyr* [Version 1.0.0; 14], *forcats* [Version 0.5.0; 9], *ggplot2* [Version 3.3.2; 10], *here* [Version 0.1; 6], *janitor* [Version 2.0.1; 4], *kableExtra* [**R-kableExtra**], *knitr* [Version 1.29; 16], *magrittr* [Version 1.5; 2], *papaja* [Version 0.1.0.9997; 1], *purrr* [Version 0.3.4; 5], *readr* [Version 1.3.1; 13], *rio* [Version 0.5.16; 3], *stringr* [Version 1.4.0; 11], *tibble* [Version 3.0.2; 7], *tidyr* [Version 1.1.0; 12], and *tidyverse* [Version 1.3.0; 15] for all our analyses.

```
## Warning: `...` is not empty.
##
```

```
## We detected these problematic arguments:
## * ‘needs_dots‘
##
## These dots only exist to allow future extensions and should be empty.
## Did you misspecify an argument?

## Warning: ‘...‘ is not empty.
##
## We detected these problematic arguments:
## * ‘needs_dots‘
##
## These dots only exist to allow future extensions and should be empty.
## Did you misspecify an argument?
```

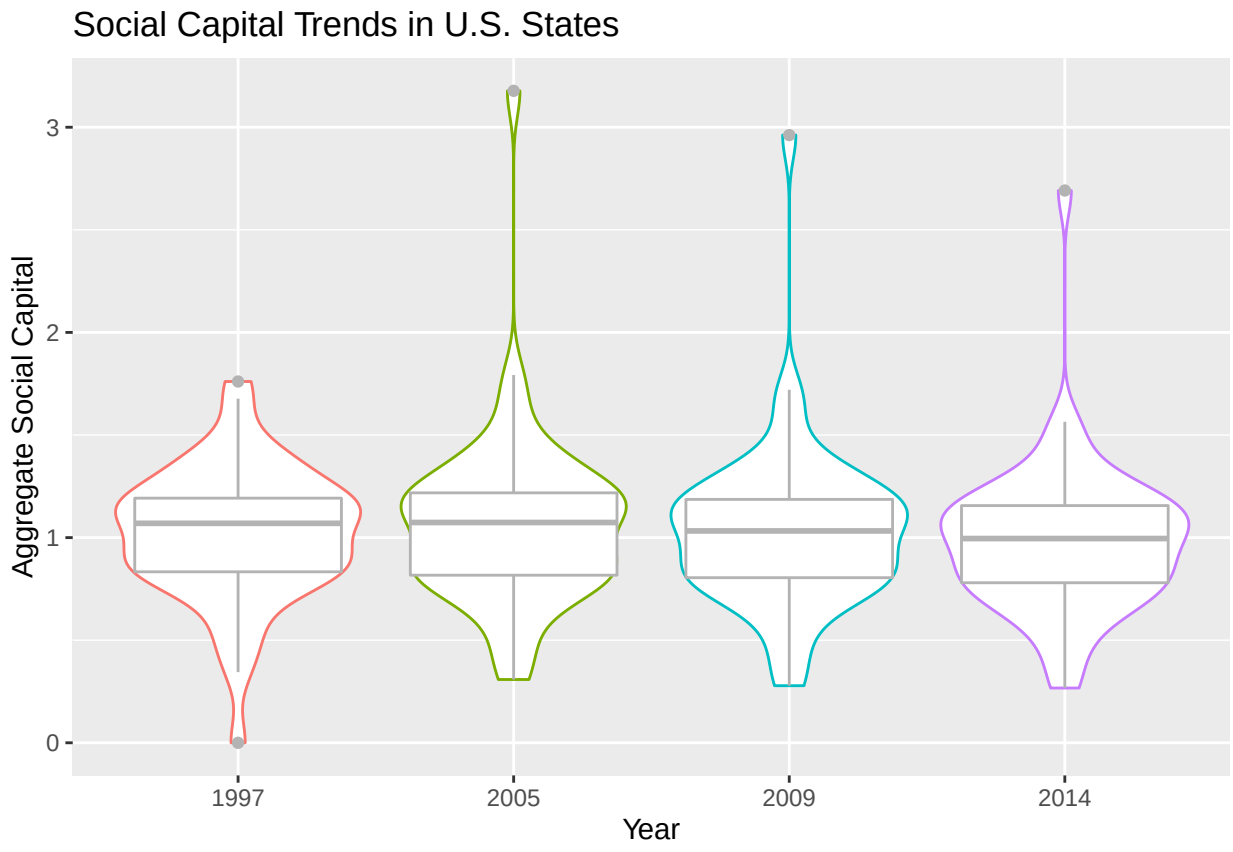
	bowl	civic	golf	relig	sport
bowl	1.00000000	0.163331564	0.1843690062	1.754235e-01	-0.011262877
civic	0.16333156	1.000000000	0.1658962995	2.547000e-01	-0.003264050
golf	0.18436901	0.165896299	1.0000000000	3.491323e-01	-0.021985208
relig	0.17542352	0.254699957	0.3491322736	1.000000e+00	0.002806421
sport	-0.01126288	-0.003264050	-0.0219852075	2.806421e-03	1.000000000
pol	-0.02966700	0.001966347	-0.0334108615	-9.423555e-05	0.012883481
prof	-0.01392921	0.076733734	-0.0417492547	-3.323368e-02	0.021631055
bus	0.09723730	0.140544743	0.1370127533	3.110755e-01	-0.020948989
labor	0.01396846	0.127298046	-0.0318697416	-4.609051e-02	0.017101150
respn	0.04573144	0.033975898	0.0008135967	-8.286148e-03	-0.004274877
pvote	0.08361274	0.095880533	0.1007224264	4.908534e-02	0.028533722
pop	-0.06101167	-0.070326839	-0.1026103152	-2.262784e-01	0.023425173

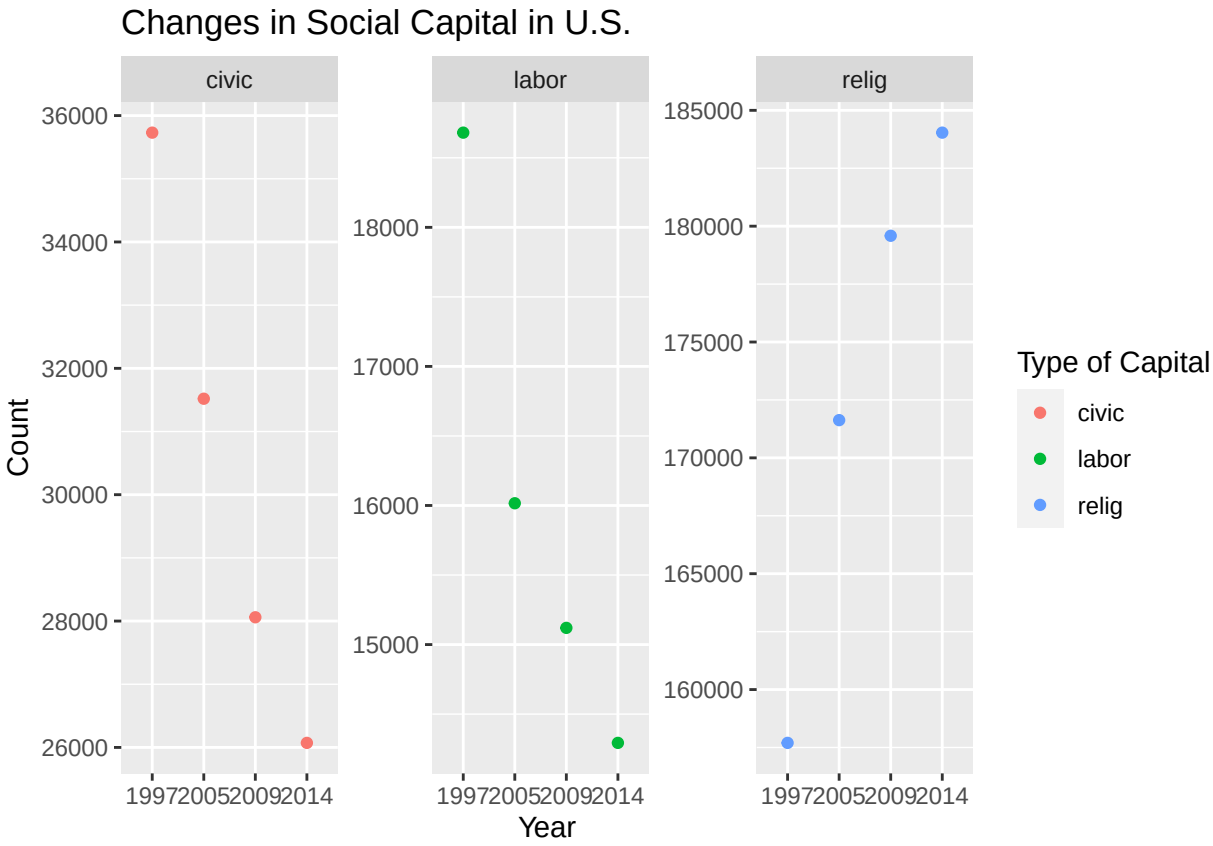
```
## nccs       0.22185326  0.347937372  0.2808439095  3.743616e-01  0.017981863
## assn       0.29451849  0.484150132  0.4950591900  9.183900e-01  0.023813113
## demmargin -0.09156467 -0.037419358 -0.1379312816 -3.337933e-01  0.024754801
##                     pol        prof         bus         labor         respn
## bowl      -2.966700e-02 -0.01392921  0.09723730  0.0139684639  0.0457314391
## civic      1.966347e-03  0.07673373  0.14054474  0.1272980463  0.0339758984
## golf      -3.341086e-02 -0.04174925  0.13701275 -0.0318697416  0.0008135967
## relig     -9.423555e-05 -0.03323368  0.31107552 -0.0460905051 -0.0082861484
## sport      1.288348e-02  0.02163105 -0.02094899  0.0171011497 -0.0042748774
## pol        1.000000e+00  0.20178507  0.08837989  0.0511035445 -0.0353395859
## prof       2.017851e-01  1.00000000  0.16314793  0.1102856692  0.0756318706
## bus        8.837989e-02  0.16314793  1.00000000 -0.0188081662 -0.1678024962
## labor      5.110354e-02  0.11028567 -0.01880817  1.0000000000  0.1920688496
## respn     -3.533959e-02  0.07563187 -0.16780250  0.1920688496  1.0000000000
## pvote      4.581321e-02  0.04152891  0.07735125 -0.0036691656  0.1131580514
## pop        4.668483e-02  0.08947146 -0.09489972  0.0586595869  0.1039281594
## nccs       8.661840e-02  0.13549486  0.32688972 -0.0002338711 -0.1209083839
## assn       6.319277e-02  0.09392249  0.47070855  0.0830621339  0.0133341334
## demmargin  8.501235e-02  0.18630848 -0.09135165  0.1315740835  0.0595954442
##                    pvote         pop          nccs        assn   demmargin
## bowl       0.083612738 -0.06101167  0.2218532641  0.29451849 -0.09156467
## civic      0.095880533 -0.07032684  0.3479373723  0.48415013 -0.03741936
## golf       0.100722426 -0.10261032  0.2808439095  0.49505919 -0.13793128
## relig      0.049085341 -0.22627840  0.3743615599  0.91838999 -0.33379330
## sport      0.028533722  0.02342517  0.0179818630  0.02381311  0.02475480
## pol        0.045813214  0.04668483  0.0866183961  0.06319277  0.08501235
## prof       0.041528912  0.08947146  0.1354948632  0.09392249  0.18630848
```

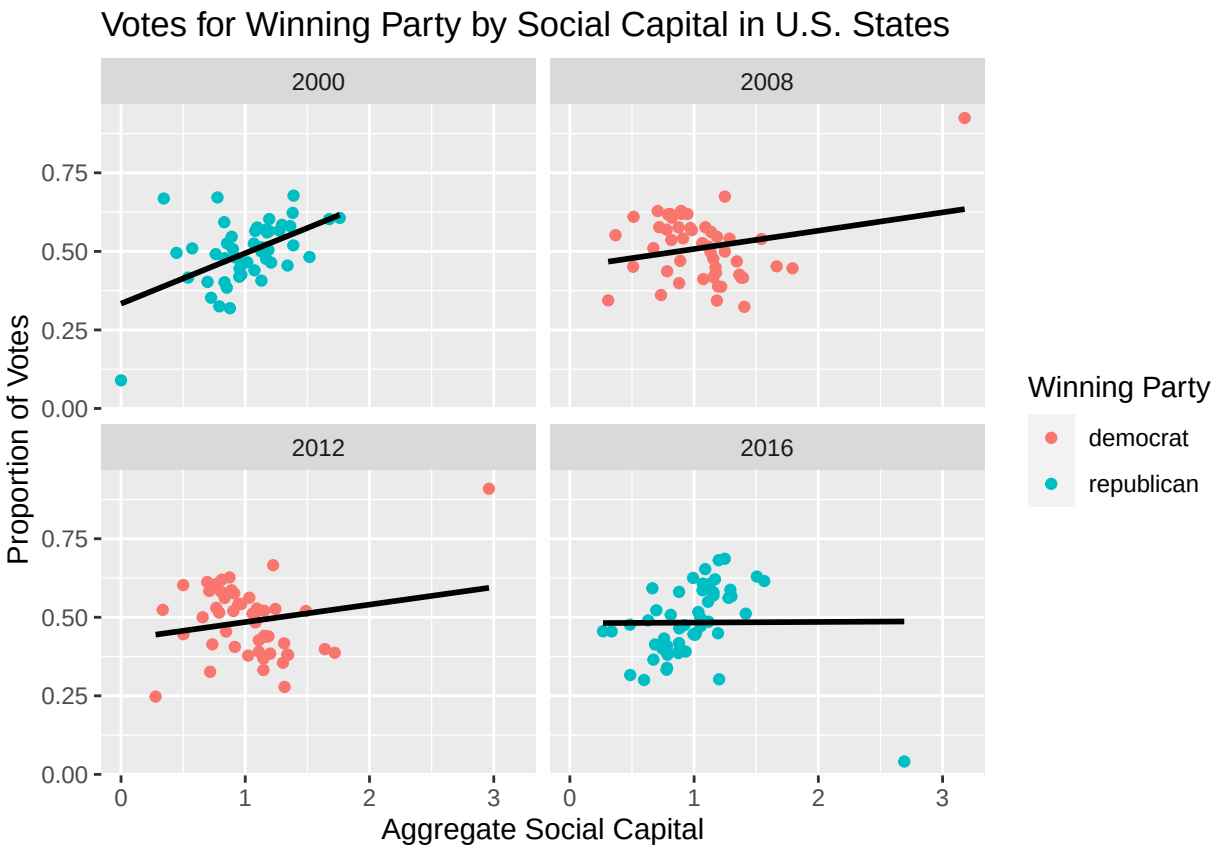
```
## bus         0.077351254 -0.09489972  0.3268897173  0.47070855 -0.09135165
## labor      -0.003669166  0.05865959 -0.0002338711  0.08306213  0.13157408
## respn       0.113158051  0.10392816 -0.1209083839  0.01333413  0.05959544
## pvote       1.000000000  0.02826091  0.2303456902  0.11604985  0.10013568
## pop         0.028260911  1.00000000 -0.0958799517 -0.19781990  0.35179190
## nccs        0.230345690 -0.09587995  1.0000000000  0.49110261 -0.06627689
## assn        0.116049854 -0.19781990  0.4911026061  1.00000000 -0.25818469
## demmargin   0.100135681  0.35179190 -0.0662768883 -0.25818469  1.00000000

##
## Call:
## lm(formula = demmargin ~ 1 + bowl + civic + golf + relig + sport +
##     pol + prof + bus + labor + pvote + respn + pop, data = df_anal_2016)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.90456 -0.18538 -0.04293  0.15323  1.21563
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.604e-01  4.802e-02  -9.588  < 2e-16 ***
## bowl        -2.034e-01  9.204e-02  -2.210   0.0272 *
## civic        6.436e-02  3.853e-02   1.671   0.0949 .
## golf        -4.419e-02  4.695e-02  -0.941   0.3467
## relig       -1.583e-01  1.110e-02 -14.253  < 2e-16 ***
## sport        1.861e-01  2.800e-01   0.665   0.5064
## pol          4.993e-01  2.240e-01   2.229   0.0259 *
## prof         1.165e+00  1.454e-01   8.011 1.60e-15 ***
```

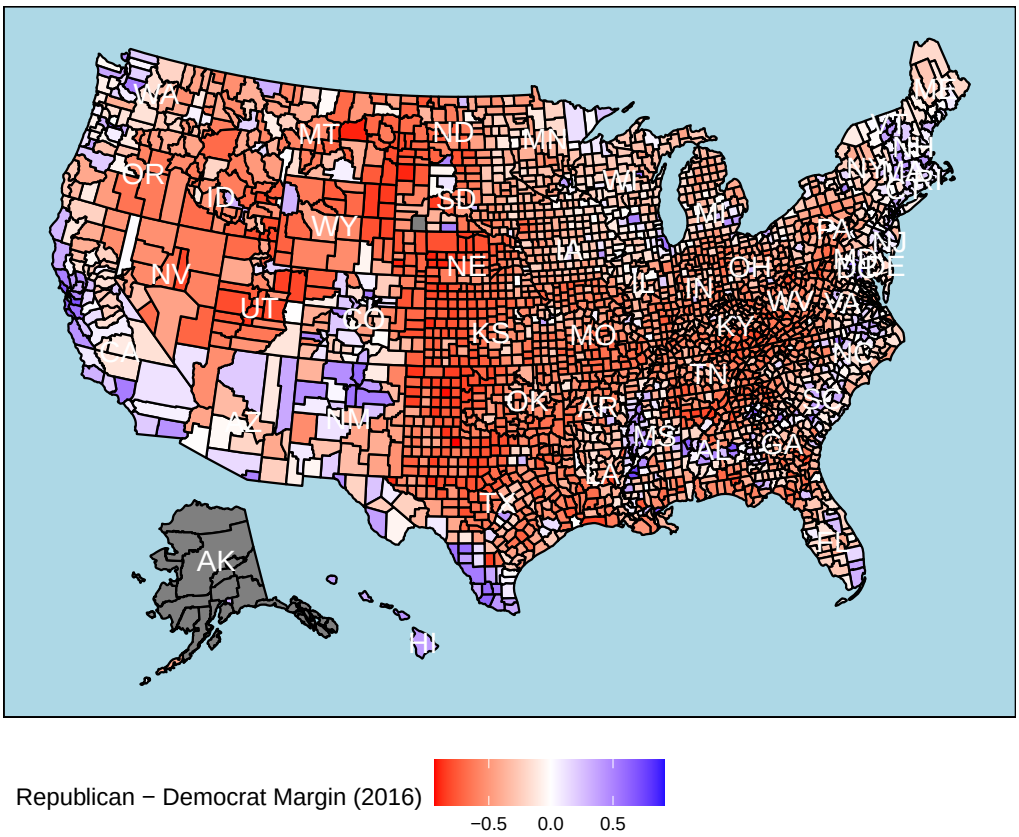
```
## bus          -4.153e-02  4.610e-02  -0.901   0.3677
## labor         4.958e-01  9.433e-02   5.256 1.58e-07 ***
## pvote         3.632e-01  5.692e-02   6.381 2.02e-10 ***
## respn        -2.638e-02  4.800e-02  -0.550   0.5826
## pop           2.653e-07  1.607e-08  16.504  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.283 on 3102 degrees of freedom
## Multiple R-squared:  0.2362, Adjusted R-squared:  0.2333
## F-statistic: 79.94 on 12 and 3102 DF,  p-value: < 2.2e-16
```

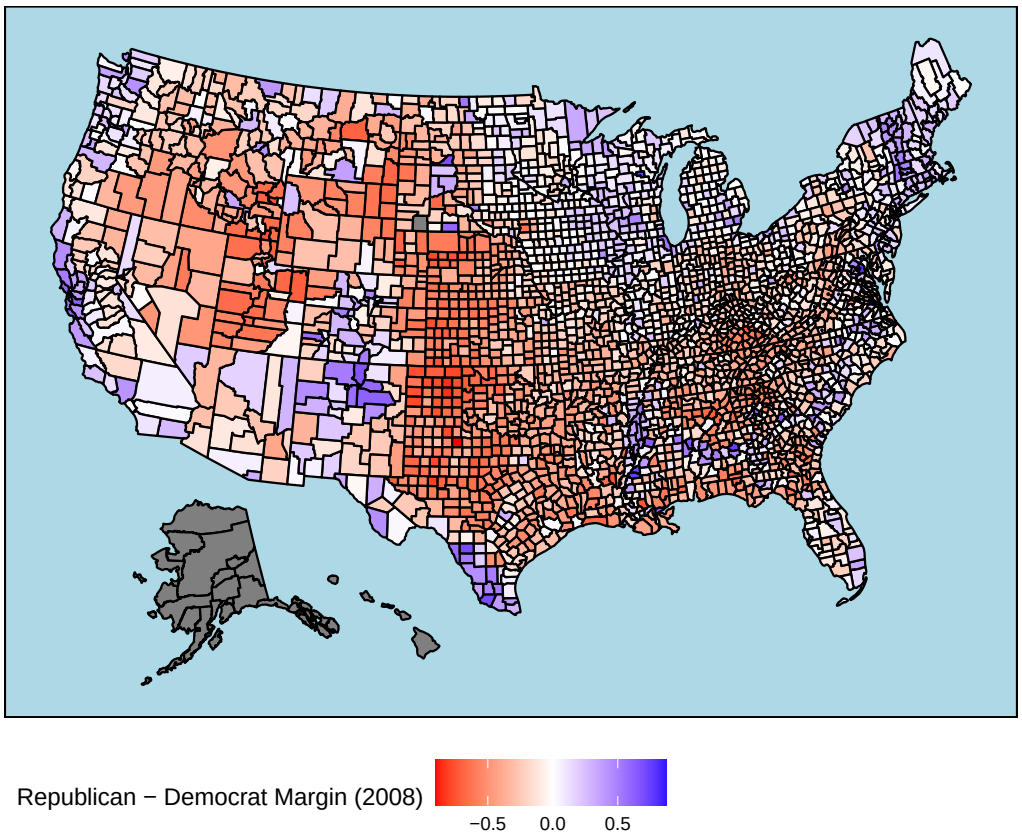



Changes in Social Capital in U.S.
civic
labor
relig
Count
Year
Type of Capital
civic
labor
relig
36000
34000
32000
30000
28000
26000
18000
17000
16000
15000
185000
180000
175000
170000
165000
160000
1997
2005
2009
2014



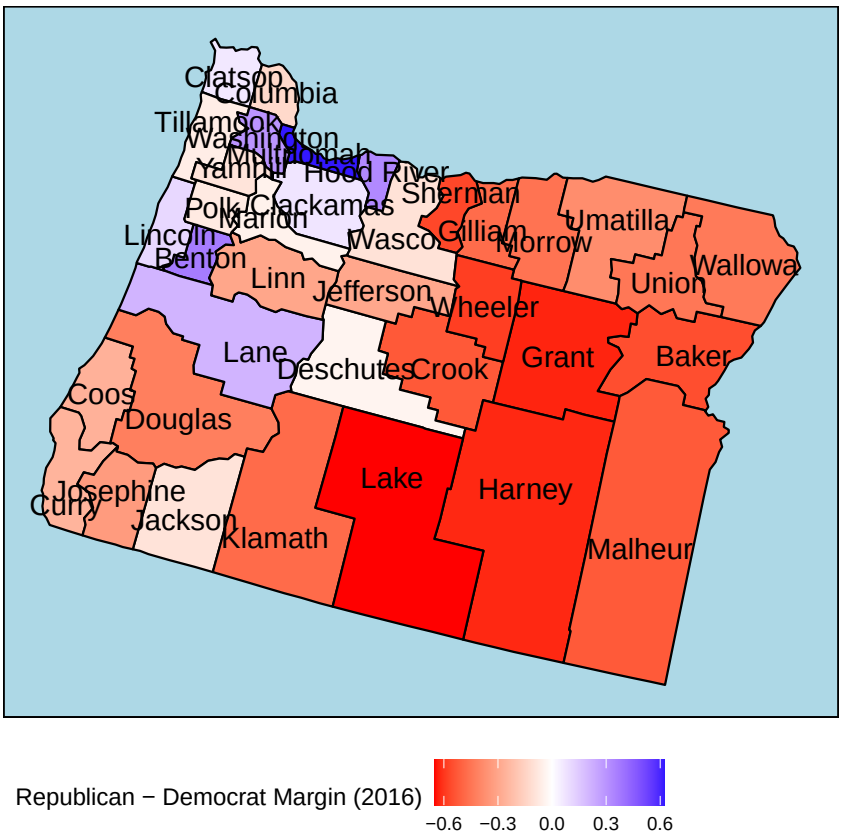

```
## Warning: package 'usmap' was built under R version 3.6.3
```

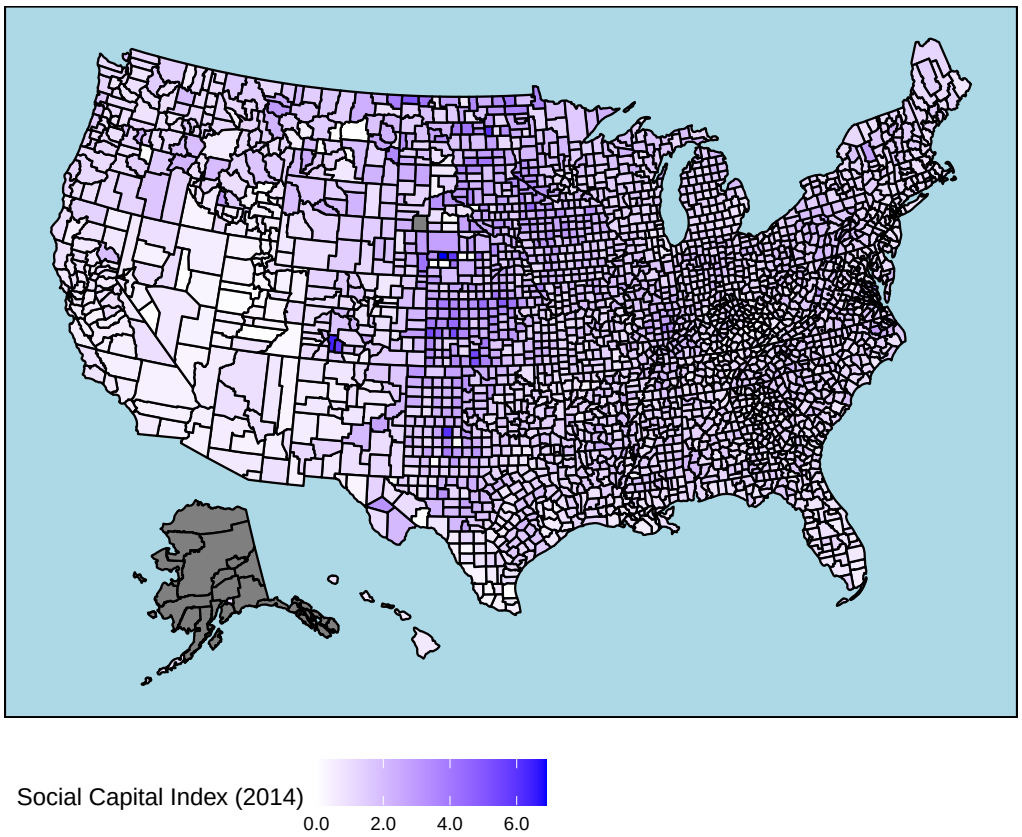

WA
MT
ND
MN
OR
ID
SD
WI
MI
WY
NY
PA
NJ
NE
IA
IL
OH
NV
UT
CO
IN
WV
VA
MD
DE
KS
MO
KY
CA
NC
TN
SC
OK
AR
AZ
NM
MS
AL
GA
LA
TX
FL
AK
HI
Republican − Democrat Margin (2016)
−0.5
0.0
0.5

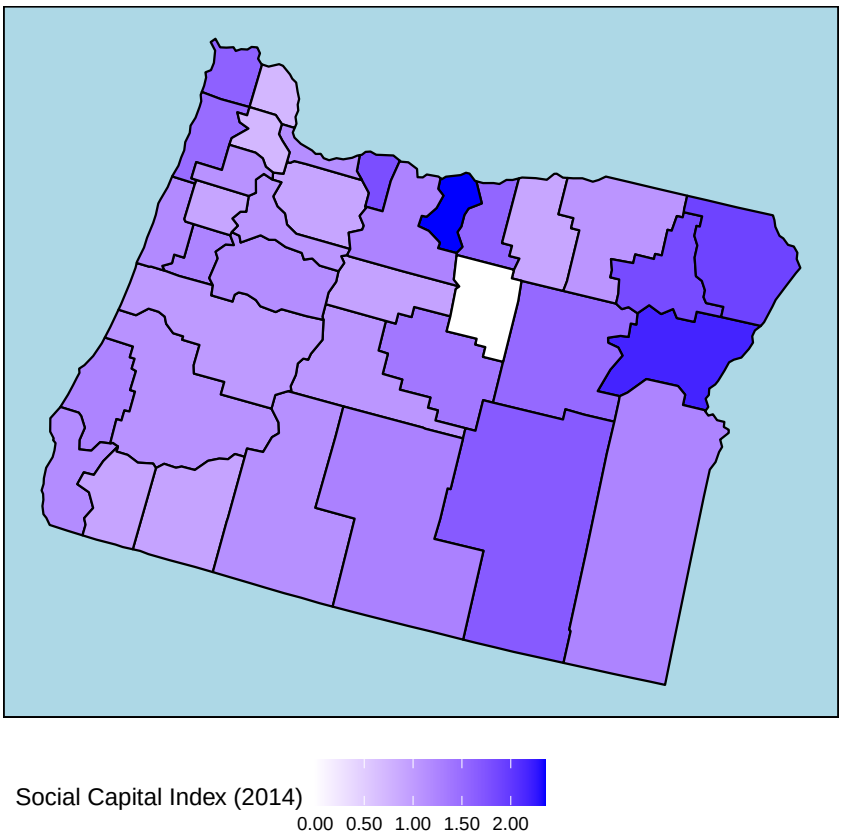



```
## [1] 2000 2008 2012 2016
```


2000
2008
2012
2016
NA
Republican − Democrat Margin (2016)
−0.5
0.0
0.5


Clatsop
Columbia
Tillamook
Washington
Multnomah
Yamhill
Hood River
Sherman
Polk
Clackamas
Marion
Lincoln
Wasco
Gilliam
Morrow
Umatilla
Benton
Linn
Jefferson
Wallowa
Union
Wheeler
Lane
Deschutes
Crook
Grant
Baker
Coos
Douglas
Lake
Harney
Josephine
Curry
Jackson
Klamath
Malheur
Republican − Democrat Margin (2016)
−0.6 −0.3 0.0 0.3 0.6


Social Capital Index (2014)
0.0
2.0
4.0
6.0



Results

Discussion

References

CSLReferences

References

[1] Frederik Aust and Marius Barth. *papaja: Create APA manuscripts with R Markdown*. R package version 0.1.0.9997. 2020. URL: https://github.com/crsh/papaja.

[2] Stefan Milton Bache and Hadley Wickham. *magrittr: A Forward-Pipe Operator for R*. R package version 1.5. 2014. URL: https://CRAN.R-project.org/package=magrittr.

[3] Chung-hong Chan et al. *rio: A Swiss-army knife for data file I/O*. R package version 0.5.16. 2018.

[4] Sam Firke. *janitor: Simple Tools for Examining and Cleaning Dirty Data*. R package version 2.0.1. 2020. URL: https://CRAN.R-project.org/package=janitor.

[5] Lionel Henry and Hadley Wickham. *purrr: Functional Programming Tools*. R package version 0.3.4. 2020. URL: https://CRAN.R-project.org/package=purrr.

[6] Kirill Müller. *here: A Simpler Way to Find Your Files*. R package version 0.1. 2017. URL: https://CRAN.R-project.org/package=here.

[7] Kirill Müller and Hadley Wickham. *tibble: Simple Data Frames*. R package version 3.0.3. 2020. URL: https://CRAN.R-project.org/package=tibble.

[8] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020. URL: https://www.R-project.org/.

[9] Hadley Wickham. *forcats: Tools for Working with Categorical Variables (Factors)*. R package version 0.5.0. 2020. URL: https://CRAN.R-project.org/package=forcats.

[10] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN: 978-3-319-24277-4. URL: https://ggplot2.tidyverse.org.

[11] Hadley Wickham. *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.4.0. 2019. URL: https://CRAN.R-project.org/package=stringr.

[12] Hadley Wickham. *tidyr: Tidy Messy Data*. R package version 1.1.2. 2020. URL: https://CRAN.R-project.org/package=tidyr.

[13] Hadley Wickham, Jim Hester, and Romain Francois. *readr: Read Rectangular Text Data*. R package version 1.3.1. 2018. URL: https://CRAN.R-project.org/package=readr.

[14] Hadley Wickham et al. *dplyr: A Grammar of Data Manipulation*. R package version 1.0.2. 2020. URL: https://CRAN.R-project.org/package=dplyr.

[15] Hadley Wickham et al. "Welcome to the tidyverse". In: *Journal of Open Source Software* 4.43 (2019), p. 1686. DOI: 10.21105/joss.01686.

[16] Yihui Xie. *Dynamic Documents with R and knitr*. 2nd. ISBN 978-1498716963. Boca Raton, Florida: Chapman and Hall/CRC, 2015. URL: https://yihui.org/knitr/.

Table 1

*(#tab:descriptives tables)Table 1
 A summary table for total votes and population by state.*

		Total Votes		Population	
Sate	N	M	SD	M	SD
AK	9	7036	399	103142	147915
AL	871	29358	44897	68880	100915
AR	975	13892	21091	37270	53303
AZ	195	142569	322062	388108	871560
CA	754	220437	459793	616298	1378441
CO	822	36534	73810	73782	148933
CT	104	195334	149300	436423	343039
DC	13	263095	45450	594503	38658
DE	39	131235	82538	282679	173390
FL	867	115569	169852	255262	407373
GA	2067	22355	50236	56445	119008
HI	12	107234	108643	355042	387067
IA	1287	15010	25994	30214	50322
ID	572	13985	27273	32809	60062
IL	1326	51236	206027	123386	528245
IN	1196	27700	45113	68419	113414
KS	1365	11003	31922	26361	70262
KY	1560	14628	32859	34967	71505
LA	832	30059	40531	70716	97338
MA	182	216390	187781	461167	393704
MD	312	104048	124062	232448	279564
ME	208	44053	39268	81514	69147
MI	1079	56095	121993	119410	266776
MN	1131	31945	79109	58910	140710
MO	1405	22511	56795	59415	120811

Table 2

*(#tab:descriptives tables)Table 2
 A summary table for votes by candidate and year of election.*

Year	Party	N	Mean Candidate Votes	SD Candidate Votes
2000	democrat	3107	16218	57150
2000	green	3107	NA	NA
2000	republican	3107	16049	38632
2000	NA	3107	339	954
2008	democrat	3108	22157	76972
2008	republican	3108	19167	44840
2008	NA	3108	577	1848
2012	democrat	3108	20974	73998
2012	republican	3108	19409	44596
2012	NA	3108	838	2952
2016	democrat	3115	21071	80496
2016	republican	3115	20160	43157
2016	NA	3115	2449	7509

Note: N = total number of counties in the US reporting data.