

Milestone #6

Debi Gonzalez

4/5/2020

Contents

1	Introduction	2
2	Overview of Original Paper & Literature Review	2
3	Proposed Extension	4
4	A Summary Table of my Data	5
5	A Beautiful Graphic	6
6	Appendix	7
	References	14

1 Introduction

Research Paper: Local demographic changes and US presidential voting, 2012 to 2016

Authors: Seth J. Hilla, Daniel J. Hopkins, and Gregory A. Huberc

This is a replication and extension project of the “Local demographic changes and US presidential voting, 2012 to 2016” study by Seth J. Hilla, Daniel J. Hopkins, and Gregory A. Huberc, which explored how demographic changes from 2000 to 2016 affected the 2016 general election outcome. The topics of immigration and the increase of racial and ethnic minorities have been fairly prevalent in media, social media, and political rhetoric over the last few years, but do these phenomena actually influence voting behavior? This study analyzes voting patterns in areas that received an influx of immigrants prior to the 2016 general election. This document presents the working replication exercise of the afore-mentioned academic paper through a textual overview of the authors’ findings and methods, a high-level visual presentation of the data, and a close-replication of the original analysis. It also contains an extension proposal, which will become the core of this project, as a form of critique, development, and/or expansion of the original findings.

Dan Hopkins offers a link to the data for this research paper on his website. The full paper can be found here. For more information about this project please visit my project’s Github repo¹.

2 Overview of Original Paper & Literature Review

As Dan Hopkins and colleagues explain in their paper, “immigration and demographic change have become highly salient in American politics, partly because of the 2016 campaign of Donald Trump.” President Trump’s victory seemed to solidify the previously proposed notion that demographic changes lead to “threatened response” in native voters realigning voting patterns on “the basis of ethnicity, nativity, nationalism, and education.” The theory implies that as demographic changes occur, native voters shift towards more populist and outspoken anti-immigration candidates. Research of anti-immigrant and Brexit, U.K. independent party, support in the United Kingdom found that these attitudes are higher in “localities that have low immigrant shares but recent demographic changes.” This trend is also true of other countries in continental Europe. In the United States, previous research with county-level data have shown that pro-GOP shifts are only associated with increases in low-skilled immigrants while increases in high-skill immigrants actually increases support for the Democratic parties. [insert source ref 27] Other studies have found an association between the percentage change in Hispanic population and shifts to the Republican Party from 2012-2016 at county-level. [insert ref13] Despite these results, the theory remains inconclusive as another study [insert ref 28] fails to find such a strong relationship between proportion change of Hispanics and Republican support using survey data and [insert ref 29] “show that the relationship between local demographic change and Trump favorability among Republicans was time-dependent.”

Notably, using precinct-level dataset of election results and demographic measures for almost 32,000 precincts in the states of Florida, Georgia, Michigan, Nevada, Ohio, Pennsylvania, and Washington, the authors of our paper of interest find that the influxes in Hispanic and foreign born groups did not influence voting behavior in favor of Trump, but rather (in a slight way) benefited his opponent in those particular areas. In other words, local demographic changes are not, on their own, increasing support for anti-immigration candidates. The authors argue that this means the precincts in question are not engaging in the often suggested “threatened response” associated with voting behavior in light of immigration influxes. The authors acknowledge that “despite its disparate local impacts, immigration may be a symbolic, nationalized issue whose effects, do not depend on local experiences.” However, the actual connection has yet to be proven empirically and it is tough to do so given several factors influencing voting patterns. Some of these factors include, but are not limited to, greater exposure to international trade and declining economic prospects for the less educated. The authors explain that ‘positive intergroup contact’ could be responsible for the seemingly ‘supportive’ voting patterns at low levels of aggregation (locally). This idea is partially supported by Enos & colleagues,

¹All analysis for this paper is available here.

whose study shows how individuals with long(er) exposures to minority groups, tend to have less exclusionary attitudes than those who are exposed to them only briefly or those who have never been exposed at all.

The authors employ regression analyses with varying explanatory variable(s) specifications and measures of ‘demographic change’ resulting in 8 main regression models as well as loess lines that visually and mathematically demonstrate the relevant results. They study change in proportion of Hispanics and proportion change of Hispanic proportions concluding that in places with demographic shifts, Trump did not benefit, but, in fact, Clinton did. These findings are further validated by a similar model utilizing foreign-born, noncitizens proportions to discard the possibility of increased Clinton support due to changing electoral composition (given that noncitizens cannot vote and would not influence the electoral outcome in such way) (Hilla, Hopkins, and Huberc 2019).

3 Proposed Extension

- 1) Examine the results at higher level of aggregation: construct same models with subset of data by state
 - I am interested in exploring this topic further by analyzing the variables used and figuring out whether this conclusion stands across geographic area, when the data is analyzed at the county or state-level. I hope to be able to do this by building the same models utilizing only a subset of the data - i.e. by state. Perhaps I will do this for only some of the states. Another way I can approach this is by grouping by county, and inputting the average values of each variable by county into the model. This would produce values relevant to the association between the explanatory variables and the outcome (gop_vote_share) at the county level.
- 2) How economic indicators affect Republican vote share: construct a model isolating economic indicators
 - I would also like to evaluate the association between economic indicators and changes in Republican voters. I have considered exploring these association in two ways: 1) Construct a model that explains shifts in GOP share of voters in terms of changes in hispanic proportion, changes in unemployment, and proportion change of poor to consider the economic influences on Republican voters share shifts. 2) Construct a model using only changes in unemployment and proportion change of poor as explanatory variables to GOP share.
- 3) Examine whether there has been shifting from majority rep/dem with changes in hispanic population.
 - To examine whether there has been majority shifts towards Republican or Democrats with changes in hispanic population, I can construct a binary variable that assigns '1' who went from less than 50% republican to more than 50% votes and another variable that does the same for Democrats. These variables can be evaluated in a regression model with hispanic proportion changes as one of the explanatory variables.

This paper evokes an interest in understanding how Trump's nationalist rhetoric and media's portrayal of immigration influxes may have attracted different demographic groups in cities that did not experience such influx (fearing that they might) i.e. the Deep South. At this point, the data used in the original paper cannot be used to explore this point as the 32,000 precincts in the sample are from the areas that have received an influx of immigrants (or experienced a demographic shift). However, this may be an interesting point to keep in mind for future research.

4 A Summary Table of my Data

Average Changes in Proportion of Population Groups by State
Proportion of foreign born and GOP voters respectively between 2011-2016

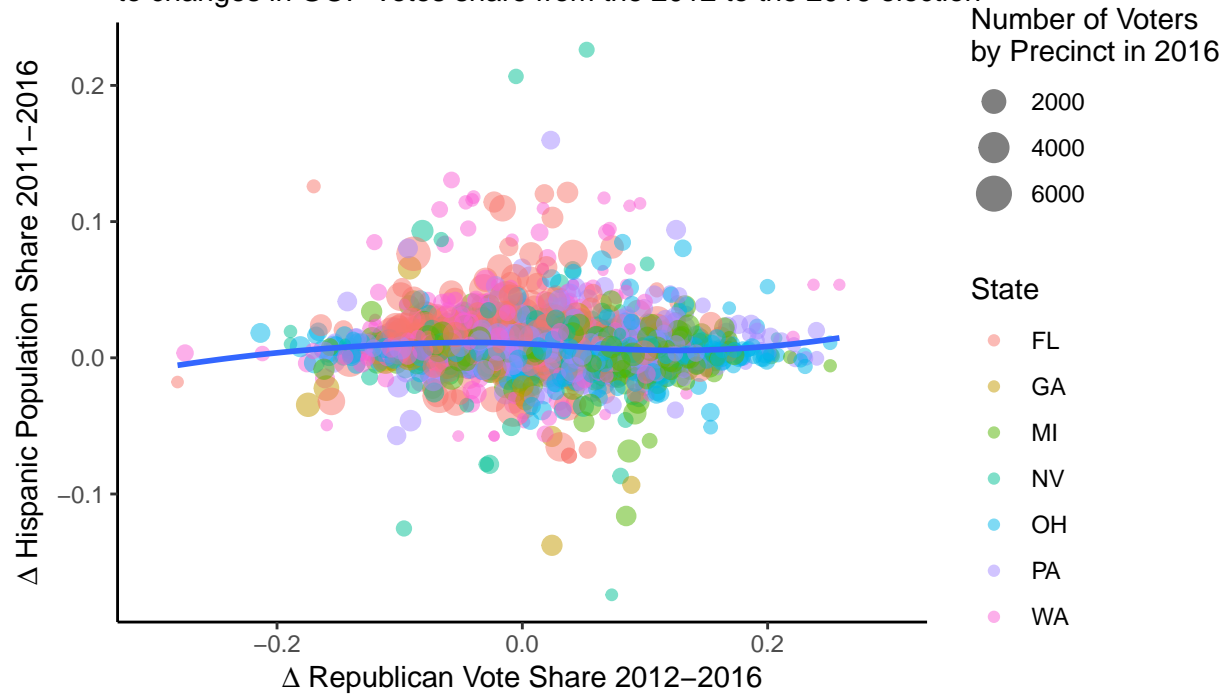
State	Avg. Change in Foreign Pop.	Avg. Change in Republican Voters
FL	-0.780	-0.162
GA	NA	NA
MI	-0.012	NA
NV	NA	NA
OH	NA	3.963
PA	0.213	2.809
WA	NA	NA

Source: HHH Demographic Threat Data Archive: Precinct

5 A Beautiful Graphic

Republican Vote Share Change vs. Hispanic Population Share Change

How changes in hispanic population share from 2011 to 2016 relate to changes in GOP Votes share from the 2012 to the 2016 election



Source: HHH Demographic Threat Data Archive: Geo_Scope

My graphic shows the association between changes of Republican vote share and Hispanic population proportion changes over the same period of time. I am using the sample function to get a random sample of 2000 observations from my geo_scope dataset. I used the geoscope dataset because I wanted to color the observations by state, which the authors do not do, but I thought would be interesting and prettier. I was able to create the loess line using geom_smooth and it does not appear dissimilar to the original paper's model(Consulting 2020). I also added the aesthetic enhancement of size as number of voters in 2016 in each precinct(Wickham et al. 2019).

To create this graphic I utilized several online guides cited in the References section of this paper(Xie, Allaire, and Grolemond 2019).

6 Appendix

6.0.0.1 Replication of Table 1: Table 1 presents multiple least-squares regression estimates of how changes in Hispanic populations correlate with increases in Republican precinct-level vote share between 2012 and 2016.

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Apr 14, 2020 - 11:52:39 PM

Table 1. Change in Republican vote share 2012 to 2016 and change in Hispanic population, various time intervals

	<i>Dependent variable:</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change in Prop. Hispanic, 2011 to 2016	−0.031*** (0.010)	−0.045*** (0.007)	−0.065*** (0.006)					
Prop. Change in Prop. Hispanic, 2011 to 2016				−0.004*** (0.0005)				
Prop. Hispanic 2011		−0.110*** (0.002)	−0.132*** (0.002)	−0.133*** (0.002)				
Prop. Change in Prop. Hispanic, 2000 to 2016								−0.005*** (0.001)
Change in Prop. Hispanic, 2000 to 2016					−0.056*** (0.007)	−0.044*** (0.005)	−0.089*** (0.005)	
Prop. Hispanic 2000						−0.122*** (0.003)	−0.132*** (0.002)	−0.147*** (0.002)
Constant	−0.00002 (0.0004)	0.055*** (0.002)	0.156*** (0.003)	0.156*** (0.003)	0.002*** (0.0005)	0.065*** (0.001)	0.132*** (0.002)	0.129*** (0.002)
Observations	28,934	28,934	28,934	28,934	28,934	28,934	28,934	28,934
R ²	0.0003	0.587	0.636	0.635	0.002	0.571	0.614	0.612
Adjusted R ²	0.0003	0.586	0.636	0.635	0.002	0.571	0.614	0.611

Note: *p<0.1; **p<0.05; ***p<0.01

I need to work on getting rid of the dependent variable row and adding the note. I also need to reorder the variable output to best match the original image.

6.0.1 Original Table

Table 1. Change in Republican vote share 2012 to 2016 and change in Hispanic population, various time intervals

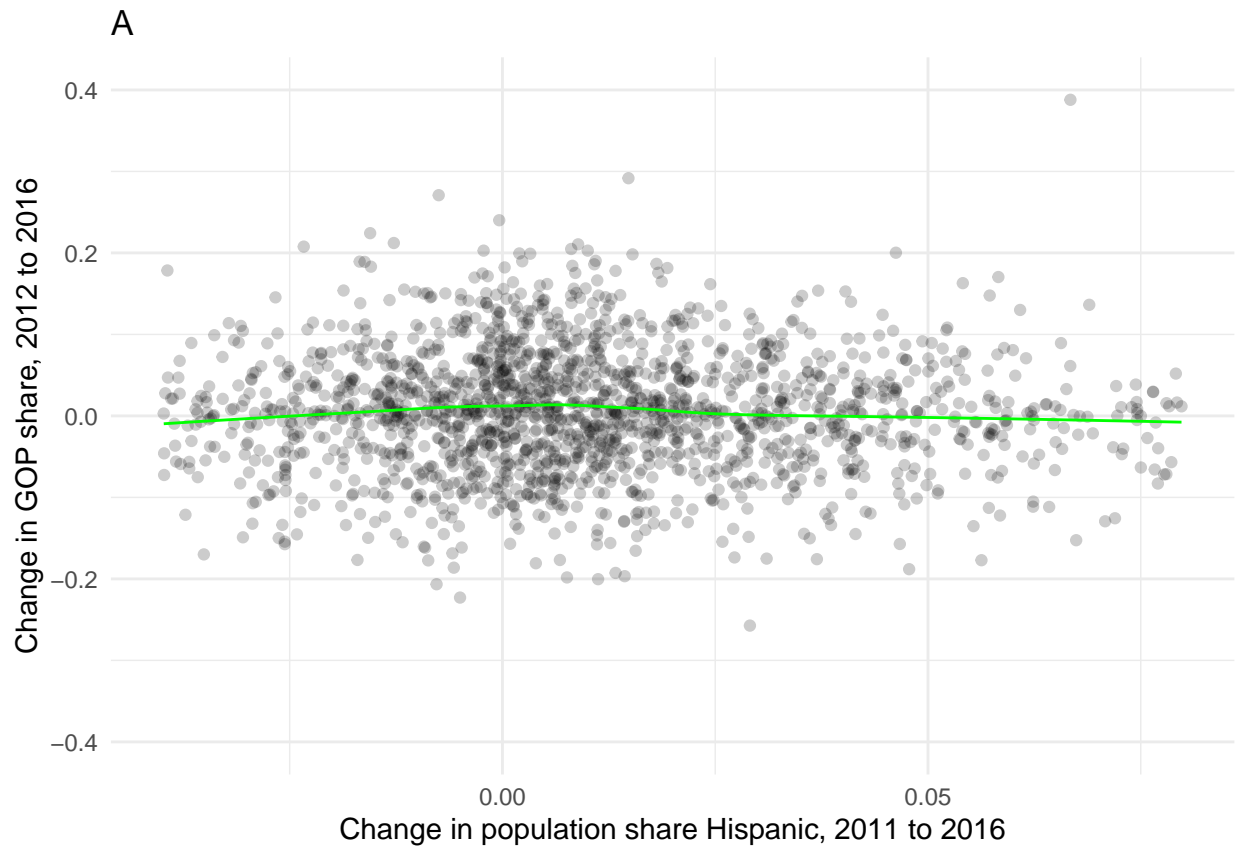
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change in prop. Hispanic, 2011 to 2016	−0.040** (0.01)	−0.071** (0.01)	−0.077** (0.01)					
Prop. Hispanic 2011		−0.13** (0.00)	−0.15** (0.00)	−0.15** (0.00)				
Prop. change in prop. Hispanic, 2011 to 2016				−0.0041** (0.00)				
Change in prop. Hispanic, 2000 to 2016					−0.077** (0.01)	−0.047** (0.01)	−0.085** (0.01)	
Prop. Hispanic 2000						−0.13** (0.00)	−0.14** (0.00)	−0.15** (0.00)
Prop. change in prop. Hispanic, 2000 to 2016								−0.0055** (0.00)
Observations	31,949	31,352	31,352	31,352	31,949	31,949	31,949	31,949
R-squared	0.001	0.658	0.704	0.704	0.004	0.649	0.689	0.687
Control for levels	No	Yes	Yes	Yes	No	Yes	Yes	Yes
County fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Additional Census controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Republican share 2012	No	No	Yes	Yes	No	No	Yes	Yes

Robust standard errors are in parentheses. * $P < 0.05$; ** $P < 0.01$. Precinct-level analysis; weighted to number of votes 2012; proportional changes top and bottom coded at 1 and −1. Note: Dependent variable is change in GOP vote share, 2012 to 2016. Prop., proportion.

Figure 1: This table offers a summary of the 8 main regression models for the original paper while omitting several control variables in the output.

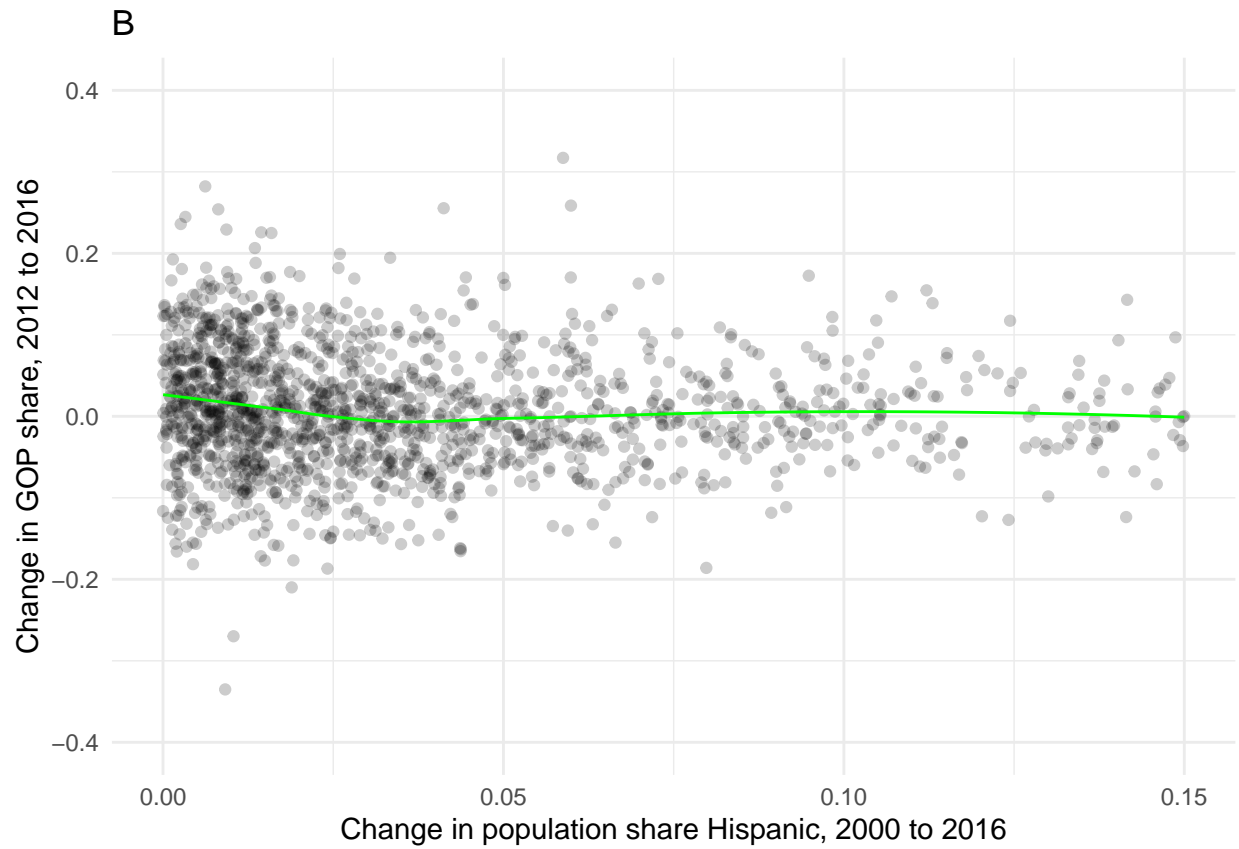
6.0.2 Figure 1

Figure 1 (A, B, C, & D) examine(s) how changes in Hispanic populations correlate with increases in Republican precinct-level vote share between 2012 and 2016. The figure shows change in the Republican share of the 2-party vote from 2012 to 2016 (positive values indicate pro-Republican shifts) against 4 different measures of change in the Hispanic population on the x axis. The 1st frame (A) measures changing population as the change in the Hispanic proportion of the overall population from 2011 to 2016, the 2nd (B) as the same change from 2000 to 2016, and the 3rd (C) and 4th (D) as proportional changes in the Hispanic population for each period.



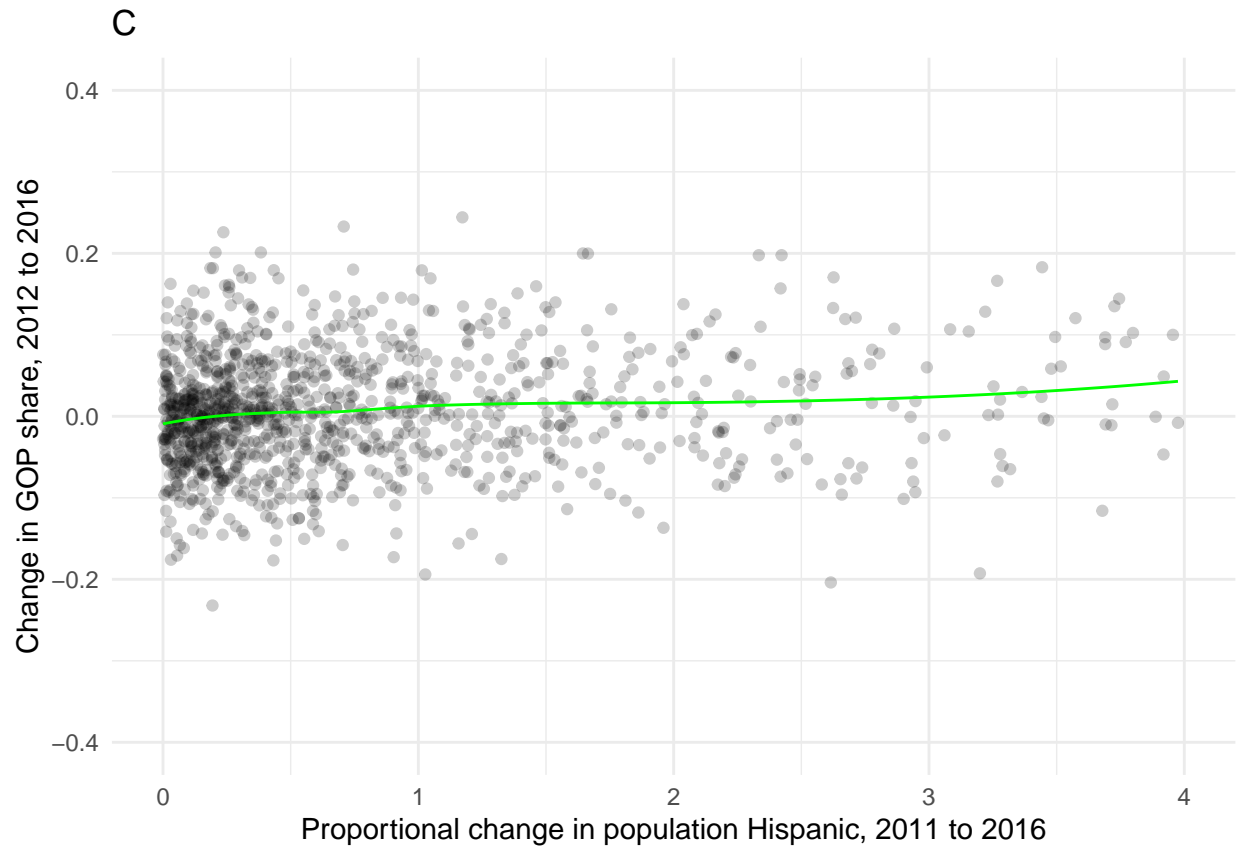
```
## Warning: Removed 423 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 423 rows containing missing values (geom_point).
```



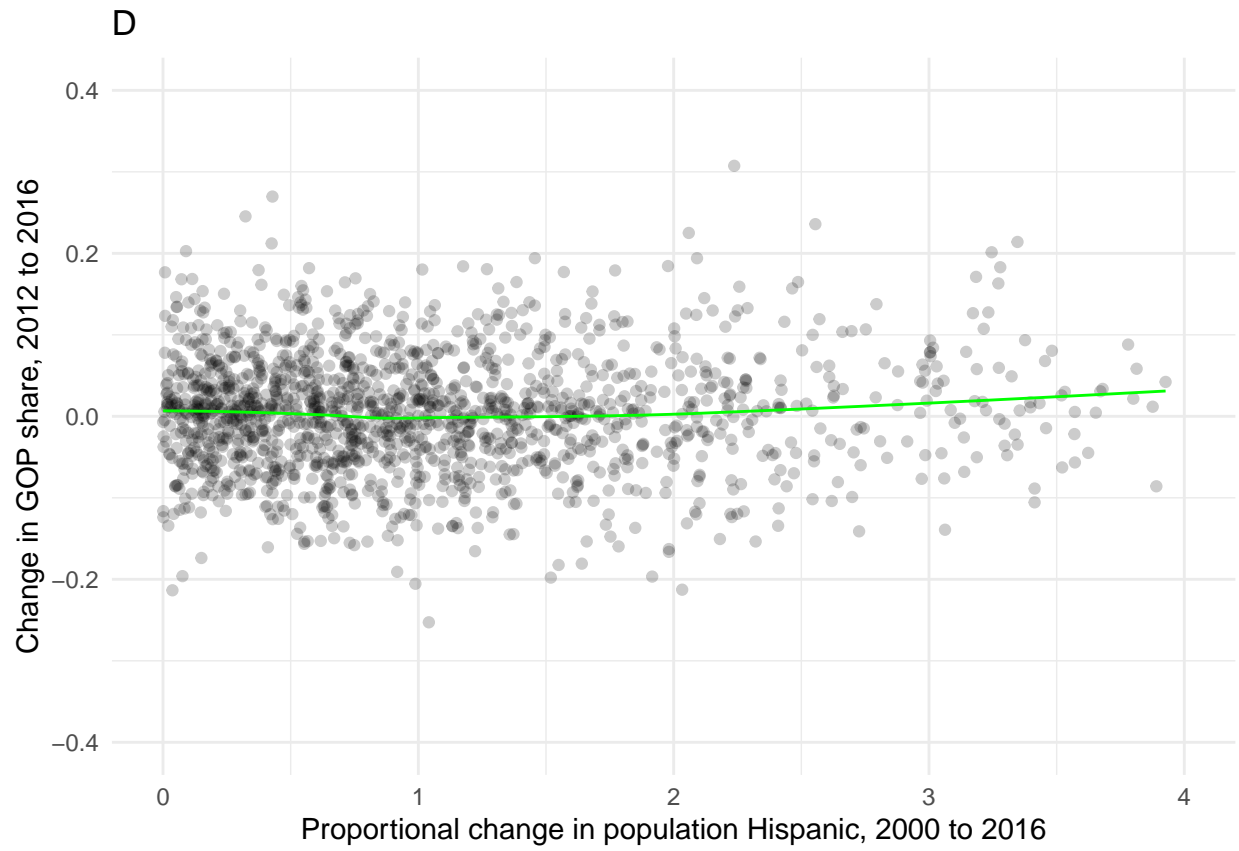
```
## Warning: Removed 859 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 859 rows containing missing values (geom_point).
```



```
## Warning: Removed 425 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 425 rows containing missing values (geom_point).
```



6.0.3 Figure 2

I am currently working to replicate figure 2, but I'm at a bit of a crossroads with the original code. Hope to update soon.

References

- Consulting, UCLA Statistical. 2020. “HOW Can I Explore Different Smooths in Ggplot2?” Available at <https://stats.idre.ucla.edu/r/faq/how-can-i-explore-different-smooths-in-ggplot2/>.
- Hilla, Seth J., Daniel J. Hopkins, and Gregory A. Huber. 2019. “Local Demographic Changes and Us Presidential Voting, 2012 to 2016.” PNAS.org Political Science.
- Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2019. “Smoothed Conditional Means.” Available at https://ggplot2.tidyverse.org/reference/geom_smooth.html.
- Xie, Yihui, J. J. Allaire, and Garrett Golemund. 2019. *R Markdown: The Definitive Guide*.