
The Rural-Urban Divide and Belief in 'America First': A Logistic Regression Analysis

Abstract

In the last election, Donald Trump rallied tremendous support from rural voters. In his campaign, Trump frequently spouted the phrase "America First." Data was gathered by the Grinnell College National Poll which surveyed 1,002 individuals asking about their behavior, identification, and opinions related to topical political issues. We hypothesized that higher proportions of those from rural areas would self-identify as believers in 'America first' than those from urban areas. In order to test this effect, we conducted a logistic regression analysis controlling for other variables that have been shown to affect political ideology. The logistic regression analysis indicated significant effects of area of residence on an individual's belief in America first. Participants located in rural areas are most likely to identify as a 'believer in America first', followed by participants located in small towns, and residents located in cities. Participants located in suburbs were the the least likely to identify as a 'believer in America first', an unexpected result. This suggests that where an individual lives predicts their view of America and its relation to the world.

Background and Significance

According to a variety of political pundits, residents of rural areas voted for Donald Trump in 2016 because they did not feel valued by mainstream politicians.⁵ A common sentiment is that the political elite abandoned rural areas in favor of American cities and other countries. This belief was not entirely unreasonable as there is evidence that increased globalization has financially benefited urban areas and has diverted resources from rural populations.² People in cities tend to hold much more globalist views than their rural counterparts.¹ While the isolationist and nationalist sentiments associated with the phrase ‘America First’ have been growing steadily in recent years, Donald Trump was one of the first to use it in modern politics.³ Trump resonated more strongly in rural areas than his Republican predecessors, possibly as a result of his highly nationalistic and non-interventionist campaign.⁴ This leads us to the question: Is there a relationship between place of residence and identification as a ‘believer in America first?’

While there seems to be an intuitive connection between rural residence, support for Trump, and identity as a ‘believer in America first’, this relationship has yet to be formally tested. The goal of the analysis is to better understand the relationship between the type of area a person resides (city, suburb, small town, rural area) and their identity as a ‘believer in America first’. In order to ensure that we measured the effects of type of area of residence alone, we controlled for a number of demographic variables. Based on existing literature, we expect rural populations and residents of small towns to self-identify as believers in ‘America First’ at higher frequencies than residents of suburbs and cities.

Table 1. Variable Names and Descriptions

Name	Description
agecat	The age category the respondent
AmericaFirst	Self-identification as a believer in America first
edu_level	Education level of the respondent
gender	The gender of the respondent
income	Annual household income category
race	Race of the respondent
religion	Religion of the respondent
residence_current	Area of residence (city, suburb, town, rural area)

Data and Methods

The data used in the analysis were gathered from the Grinnell National Poll (GCNP). The GCNP data are available through the Grinnell College Data Analysis and Social Inquiry Lab. The data was collected by Ann Selzer & Co. through phone interviews of 1,002 adults from across the United States from August 29 to September 2, 2018. The GCNP asked these individuals a number of questions on issues pertaining to politics.⁶ Selzer & Co. provided weights with the raw data. When those weights are applied, the dataset becomes more representative of the overall population.

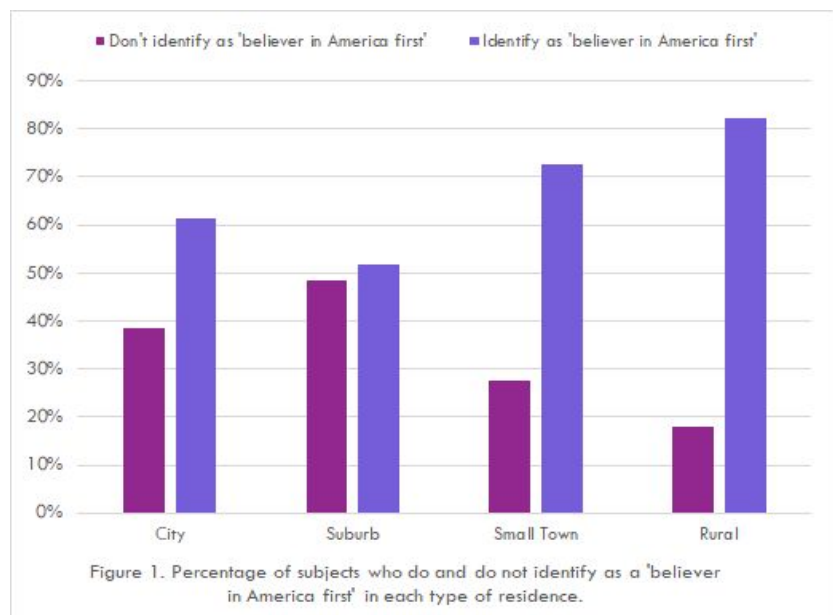
The dataset included seven variables that pertained to the question of our study. Within current residence, the explanatory variable of interest, 23 subjects fell into the “Refused/Not Sure” category and were consequently removed from the dataset. Including non-responses in the analysis would not assist in any conclusive inferences. It was not necessary to remove any subjects due to missingness for the control variables. After cleaning the data, the sample size totalled 979 observations.

The variable that measured whether or not a subject self-identified as a 'believer in America first' was originally reported in eleven separate variables depending on the order in which the subject listed their identities. We cleaned the data by collapsing columns into one column, creating a binary dummy variable called 'America First'. Current residence, the predictor of interest, and all control variables, were coded categorically. These variables were re-coded into dummy variables for the analysis.

In the analysis, a multiple logistic regression model was performed to examine the relationship between current residence and self-identification as a 'believer in America first.' The response variable was the subject's self-identification as a 'believer in America first.' The explanatory variable of interest was residence, which included four categories: city, suburb, small town, and rural. We included a number of control variables: education level, age category, religion, race, income, and gender, which reduces potential omitted variable bias (see full model in Appendix B). In specifying the model, we selected these variables as controls as they have been shown to predict political ideologies of individuals.

Results

After conducting the logistic regression, we found that city, suburb, and rural residences have significant relationships with self-identification as a 'believer in America first' ($p < .05$). In the analysis, we used city as the baseline to calculate odds ratios for the other three types of residences. Subjects located in suburbs are 1.6 times less likely to identify as a 'believer in America first' than residents of cities. While location in a small town is not significantly different from city in the model ($p=0.053$), subjects located in small towns are 1.5 times more likely to identify as a 'believer in America first' than residents of cities. Finally, subjects located in rural areas are 2.3 times more likely to identify as a 'believer in America first' than residents of cities. Refer to Appendix C for the entire regression output.



Discussion

The model appears to meet the assumptions of logistic regression. Firstly, the response variable is a binary variable. A variance inflation factors test indicated that there is no severe multicollinearity among the predictors as the values are less than 3 (Appendix D). However, the correlation between the variables may be artificially lower because they are all categorical variables. All observations are independent of each other as the data was collected through random phone surveys. Lastly, there are no outliers as the variables are all categorical and the sample size is large. The assumption that there is a linear relationship between the logit of the

response and each predictor is not applicable, as the variables are categorical. We believe all applicable assumptions are met.

The statistical analysis shows current residence and self-identification as a 'believer in America first' have a significant relationship controlling for education level, age, religion, race, income, and gender. With city as the baseline for comparison, suburb and rural residences were significant. There is a noteworthy, but non-significant relationship between small town residence and 'believer in America first' ($p = .053$).

These results show that an individual living in a rural area is 2.3 times more likely to be a 'believer in America first', than an individual living in a city. This confirms our hypothesis that those in rural areas are more likely to identify as a 'believer in America first.' Individuals living in small towns are 1.5 times more likely to identify as a 'believer in America first.' This would also support that our hypothesis is accurate. Subjects located in suburbs are 1.6 times less likely to identify as a 'believer in America First' than residents of cities, an unexpected result we would like to study in future research.

No causality can be inferred from this survey and analysis as place of residence can influence self-identification as a 'believer in America first', or self-identification as a 'believer in America first' can influence where one chooses to live. Political ideologies may naturally encourage one to live in a certain area of residence. On the other hand, one may move to an area and have their ideologies changed in order to conform to the common thought of the area of residence.

Our research was limited by some factors. The data were gathered through phone surveys. Due to the nature of phone surveys, many participants may choose not to respond, leading to potential selection bias. Surveyors attempted to overcome and control for this potential bias through weighting the data collected from respondents. This is discussed above in the data section. However, potential bias may still exist in the data.

The dataset is fairly large with 979 observations. This supports the reasonableness of the conclusions drawn from our analysis and supports the assumption that our dataset is reflective of the represented groups in the data. In our analysis, a large number of control variables were included in the model and this is justified by the large sample size.

The results of this research suggest that there are strong ties between place of residence and political ideology in American society. We did not expect that the likelihood of suburbs would be less likely to identify as a 'believer in America first' than cities. We could more closely examine the differences in political self-identification between these two areas of residence in other studies. We also suggest conducting future research to evaluate the regional differences in political self-identification. In addition, it would be interesting to see how rhetoric such as Donald Trump's references to "America First" affect voting patterns through various demographics.

References

1. Curtis, S. (2017). Global Cities at the End of Globalism: Can they survive? *The Chicago Council on Global Affairs*. Retrieved from <https://www.thechicagocouncil.org/blog/global-insight/global-cities-globalism-nationalism-populism-can-they-survive>
2. Green, G. P. (2016). Rural America in a Globalizing World: Problems and Prospects for the 2010s. *Contemporary Sociology*, 45(1), 18–20. <https://doi.org/10.1177/0094306115621526a>
3. Kagan, R. (2018, September 23). 'America First' Has Won. *The New York Times*. Retrieved from <https://www.nytimes.com/2018/09/23/opinion/trump-foreign-policy-america-first.html>
4. Kurtzleben, D. (2016, November 14). Rural Voters Played A Big Part In Helping Trump Defeat Clinton. *NPR*. Retrieved from <https://www.npr.org/2016/11/14/501737150/rural-voters-played-a-big-part-in-helping-trump-defeat-clinton>
5. Leonard, R. (2017, January 05). Why Rural America Voted for Trump. Retrieved from <https://www.nytimes.com/2017/01/05/opinion/why-rural-america-voted-for-trump.html>
6. Selzer and Company (2018). *Grinnell College National Poll* [Data file]. Retrieved from <https://dasil.sites.grinnell.edu/political-science/data-from-the-grinnell-college-national-poll>

Appendix

Appendix A: Level Names

Name in Results	Category Name
as.factor(race)1	White
as.factor(race)2	Black
as.factor(race)3	Hispanic
as.factor(race)4	Asian
as.factor(race)5	Other
as.factor(race)6	Multi-Racial
as.factor(race)7	Refused/Not Sure
as.factor(agecat)1	Under 25
as.factor(agecat)2	25 to 29
as.factor(agecat)3	30 to 34
as.factor(agecat)4	35 to 39
as.factor(agecat)5	40 to 44
as.factor(agecat)6	45 to 49
as.factor(agecat)7	50 to 54
as.factor(agecat)8	55 to 59
as.factor(agecat)9	60 to 64
as.factor(agecat)10	65 and above
as.factor(agecat)11	Refused/Not Sure
as.factor(edulevel)1	High School or less
as.factor(edulevel)2	Some college
as.factor(edulevel)3	Associates degree
as.factor(edulevel)4	Bachelor's degree
as.factor(edulevel)5	Postgraduate
as.factor(edulevel)6	Refused/Not Sure
as.factor(residence_current)1	City
as.factor(residence_current)2	Suburb
as.factor(residence_current)3	Small Town
as.factor(residence_current)4	Rural
as.factor(religion)1	Roman Catholic
as.factor(religion)2	Protestant
as.factor(religion)3	Some other kind of Christian
as.factor(religion)4	Jewish
as.factor(religion)5	Muslim
as.factor(religion)6	Other
as.factor(religion)7	None
as.factor(religion)8	Refused/Not Sure
as.factor(income)1	Under \$25,000
as.factor(income)2	\$25,000-\$49,999
as.factor(income)3	\$50,000-\$74,999
as.factor(income)4	\$75,000-\$99,999
as.factor(income)5	\$100,000 or more
as.factor(income)6	Refused/Not Sure
as.factor(gender)1	Male
as.factor(gender)2	Female
as.factor(gender)3	Other

Appendix B: Full Model

Model = $\beta_0 + \beta_1 \cdot \text{Black} + \beta_2 \cdot \text{Hispanic} + \beta_3 \cdot \text{Asian} + \beta_4 \cdot \text{Other} + \beta_5 \cdot \text{Multi-Racial} + \beta_6 \cdot$
 Refused or Not sure (Race) + $\beta_7 \cdot 25 \text{ to } 29 + \beta_8 \cdot 30 \text{ to } 34 + \beta_9 \cdot 35 \text{ to } 39 + \beta_{10} \cdot 40 \text{ to } 44 + \beta_{11} \cdot$
 45 to 49 + $\beta_{12} \cdot 50 \text{ to } 54 + \beta_{13} \cdot 55 \text{ to } 59 + \beta_{14} \cdot 60 \text{ to } 64 + \beta_{15} \cdot 65 \text{ and over} + \beta_{16} \cdot$
 Refused or Not Sure (Age) + $\beta_{17} \cdot \text{Some college} + \beta_{18} \cdot \text{Associates} + \beta_{19} \cdot \text{Bachelor's} + \beta_{20} \cdot$
 Postgraduate + $\beta_{21} \cdot \text{Refused or Not Sure (Education)} + \beta_{22} \cdot \text{Suburb} + \beta_{23} \cdot \text{Small Town} + \beta_{24} \cdot$
 Rural + $\beta_{25} \cdot \text{Protestant} + \beta_{26} \cdot \text{Some other kind of christian} + \beta_{27} \cdot \text{Jewish} + \beta_{28} \cdot \text{Muslim} + \beta_{29} \cdot$
 Other(Religion) + $\beta_{30} \cdot \text{None} + \beta_{31} \cdot \text{Refused or Not Sure(Religion)} + \beta_{32} \cdot \$25,000 - \$49,999 +$
 $\beta_{33} \cdot \$50,000 - \$74,999 + \beta_{34} \cdot \$75,000 - \$99,999 + \beta_{35} \cdot \$100,000 \text{ or more} + \beta_{16} \cdot$
 Refused or Not Sure (Income) + $\beta_{37} \cdot \text{Female} + \beta_{38} \cdot \text{Other}$

Appendix C: Regression Output

Call:

```
glm(formula = AmericaFirst ~ as.factor(race) + as.factor(agecat) + as.factor(edulevel) +  
as.factor(residence_current) + as.factor(religion) + as.factor(income) + as.factor(gender), family  
= binomial("logit"), data = GCNP, weights = weight)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.8972	-0.8379	0.4346	0.7828	2.7282

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.43320	0.39337	1.101	0.27078
as.factor(race)2	-0.28123	0.24753	-1.136	0.25589
as.factor(race)3	-0.61472	0.26124	-2.353	0.01862 *
as.factor(race)4	0.41940	0.44696	0.938	0.34807
as.factor(race)5	-0.29381	0.44769	-0.656	0.51164
as.factor(race)6	-0.75566	0.85668	-0.882	0.37773
as.factor(race)7	-0.08948	0.45275	-0.198	0.84333
as.factor(agecat)2	-0.39425	0.35353	-1.115	0.26477
as.factor(agecat)3	0.43041	0.33000	1.304	0.19213
as.factor(agecat)4	0.66355	0.35847	1.851	0.06416 .
as.factor(agecat)5	0.76174	0.34201	2.227	0.02593 *
as.factor(agecat)6	1.19119	0.38922	3.060	0.00221 **
as.factor(agecat)7	1.20905	0.37574	3.218	0.00129 **
as.factor(agecat)8	0.80681	0.33261	2.426	0.01528 *
as.factor(agecat)9	0.21095	0.32352	0.652	0.51436
as.factor(agecat)10	0.84390	0.30265	2.788	0.00530 **
as.factor(agecat)11	0.56169	0.86719	0.648	0.51717
as.factor(edulevel)2	0.01829	0.23123	0.079	0.93695
as.factor(edulevel)3	-0.30064	0.28105	-1.070	0.28475
as.factor(edulevel)4	-0.65972	0.23543	-2.802	0.00508 **
as.factor(edulevel)5	-1.49658	0.27922	-5.360	8.33e-08 ***
as.factor(edulevel)6	-1.24475	0.92367	-1.348	0.17778
as.factor(residence_current)2	-0.45469	0.20497	-2.218	0.02653 *
as.factor(residence_current)3	0.40830	0.21125	1.933	0.05326 .
as.factor(residence_current)4	0.83304	0.25365	3.284	0.00102 **
as.factor(religion)2	0.68313	0.29542	2.312	0.02076 *
as.factor(religion)3	0.50440	0.25436	1.983	0.04736 *
as.factor(religion)4	-0.27925	0.68498	-0.408	0.68351
as.factor(religion)5	-0.50059	0.73847	-0.678	0.49785
as.factor(religion)6	-0.39019	0.27159	-1.437	0.15080
as.factor(religion)7	-0.74956	0.27113	-2.765	0.00570 **
as.factor(religion)8	-0.70302	0.61431	-1.144	0.25245
as.factor(income)2	0.28063	0.25338	1.108	0.26806
as.factor(income)3	0.10678	0.26996	0.396	0.69243
as.factor(income)4	-0.01490	0.29371	-0.051	0.95954
as.factor(income)5	0.04576	0.27087	0.169	0.86585
as.factor(income)6	0.73797	0.29789	2.477	0.01324 *
as.factor(gender)2	-0.25531	0.16125	-1.583	0.11336
as.factor(gender)3	10.19526	535.41134	0.019	0.98481

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1233.8 on 978 degrees of freedom

Residual deviance: 1037.2 on 940 degrees of freedom

AIC: 1132.7

Number of Fisher Scoring iterations: 12

Appendix D: Variance Inflation Factor (VIF) Table

Black 1.154567	Hispanic 1.389165	Asian 1.096371	Other 1.076288
Multi-Racial 1.036396	Refused/Not Sure (race) 1.156661	25 to 29 1.505967	30 to 34 1.1.764390
35 to 39 1.706654	40 to 44 1.740155	45 to 49 1.474492	50 to 54 1.510316
55 to 59 1.746216	60 to 64 1.781446	65 and over 2.102813	Some College 1.367202
Associates Degree 1.287265	Bachelor's Degree 1.569728	Postgraduate Degree 1.590515	Refused/Not Sure (Education) 1.066893
Suburb 11.578165	Small Town 1.465575	Rural 1.371860	Protestant 1.643050
Some other Christian 2.223942	Jewish 1.115378	Muslim 1.117036	Other 2.013779
None 1.995316	Refused or Not Sure (religion) 1.221835	\$25000-\$49999 1.700249	\$50000-\$74999 1.614321
\$75000-\$99999 1.643247	\$100000 or more 2.214701	Female 1.131950	Other 1.000001

Appendix E: Analysis of Deviance Table

Model: gaussian, link: identity

Response: AmericaFirst

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			1001	228.64
race	1	0.4579	1000	228.18
agecat	1	4.8844	999	223.30
edulevel	1	10.3965	998	212.90
residence_current	1	4.3109	997	208.59
religion	1	6.1683	996	202.42
income	1	0.1925	995	202.23
gender	1	0.0211	994	202.21