

WHO DECIDES THE ELECTION?

TECHNICAL REPORT

KIESKOMPAS – ELECTION COMPASS, COMMISSIONED BY THE FRIEDRICH-EBERT-STIFTUNG
OCTOBER 2020

AUTHORS: TOM ETIENNE, LAURA VAN HECK & ANDRÉ KROUWEL

CONTRIBUTORS: DAVID PUPOVAC, YORDAN KUTIYSKI, ANDREW PASQUIER

› Contents

• Methodology	3
• Extant research: data and literature study	3
• Conducted survey	3
• Targeted sampling	4
• Weighting	5
• Selection and categorization of voter groups	6
• Conditional classification	6
• Clustering	6
• Cluster analysis dimensions	6
• Cluster analysis methodology	8
• Final clustering solution	9
• Cluster segmentation and interpretation	9
• Analysis	10
• Bibliography	12

This technical report serves as supplementary material for the study 'Who Decides the Election: A Study of Decisive Voter Groups Considering a Biden Vote', October 2020.

› **Methodology**

In this technical report, we will elaborately discuss the methodological techniques applied to the ‘Who Decides the Election’ study. Specifically, we focus on how the literature and data provide a framework for the applied targeted sampling approach, followed by the structure and methods used for sampling, the selection of a voter contingent of interest, and the clustering of that contingent. A concise version of this report is included in the study under the methodology chapter.

In order to most effectively analyze psychographic and ideological opinion structures of the 2020 decisive voter contingent considering a vote for Joe Biden, we have initially examined existing academic literature, journalism, data and polling. This identification served as the basis to, in a purposeful and targeted manner, sample respondents to participate in a survey (see survey questions in the technical report). Subsequently, we grouped respondents in five broad categories, based on which candidate they were more likely to support: Trump voters (1), Trump sympathizers (2), respondents sympathizing with both Trump and Biden (3), Biden sympathizers (4) and Biden voters (5). The voter contingent considering a Biden vote (i.e. groups 3 and 4) was thereafter statistically clustered to identify the most homogenous clusters within the cohort.

› **Extant research: data and literature study**

First, we conducted a multi-faceted review of recent election data, opinion polling, political journalism, and academic literature to ground our research within ongoing debates and critical understandings of the US electorate. At a practical level, this literature- and data-based study provides a framework for the applied targeted sampling approach used for the survey. For the comparative study of past and current polling, we relied on both raw polling results from reputable sources like Pew and Gallup, as well as meta-reports and summaries from think-tanks like Brookings and the Center for American Progress. Where relevant, we also cited election coverage from media outlets that spotlighted specific voter groups, since popular perceptions about voting patterns impacts media narratives and campaign strategies, that then in turn can influence real results.

› **Conducted survey**

The survey fielded was developed with the objective of identifying crucial voter groups, focusing on instrumental explanatory factors to achieve this aim. The first survey module consisted of two blocks of questions: one with demographic items and one with political affiliation questions. For both of these blocks, we predominantly drew measures and wording from the American National Election Study (ANES, 2017) as well as other renowned surveys on US politics. The political affiliation question block contains items measuring the various ways in which US citizens express their political identification, including party identification, ideological orientation, vote recall for the last two presidential elections, vote intention in the current elections, propensity to vote questions about both current presidential candidates and the Democratic and Republican parties in general, as well as generic economic left-right and cultural progressivism-conservatism self-placement scales. The block was complemented with a set of questions asking whether a respondent’s vote was more pro-Biden or anti-Trump for those contemplating a Biden vote, or more anti-Biden versus pro-Trump for those contemplating a Trump vote.

All question blocks that follow, with the exception of the closing block, were randomized. Each respondent, provided they completed the survey, saw those blocks in random order. Most questions and statements were asked on a five-point Likert scale.

A first module of blocks maps the psychographics of respondents, asking items about authoritarianism, patriotism, and conspiracy beliefs. The set of authoritarianism questions was drawn partly from the British Social Attitudes Survey (Curtice et al., 2020). The set of questions concerning constructive patriotism and nationalism was combined with questions measuring opinions about US American exceptionalism. For questions about the first two concepts, we drew inspiration from the ISSP (Smith, 1992) and Davidov’s analysis of measurement equivalence nationalism and constructive patriotism (2009). The questions mapping conspiracy mentalities are taken from Bruder et al.’s Conspiracy Mentality Questionnaire (2013).

A second module of question blocks asked about economic evaluations. We asked respondents to consider their relative deprivation, mapped respondent precarity, and asked about convictions regarding the role of government and economic intervention. Perceived relative deprivation, the feeling of being left behind whilst others are allowed to profit (Walker & Smith, 2002), has recently more prominently been linked with populist voting behavior. Hameleers (2020) shows that the effects of populist messages on populist attitudes are contingent, among others, on relative deprivation. Like Hameleers, we drew from Elchardus's (2014) study on populism, persistent republicanism and declinism. The questions about precarity are a selection of Antonucci's (2017) study on the squeezed middle. The questions about the role of government stem from the 2007 NILT survey (Devine, 2007). The economic position statements are drafted largely by Kieskompas - Election Compass experts.

A third module consists of question blocks concerning political attitudes, leadership evaluations, campaign evaluations and institutional trust. This module contains items specifically pertaining to the current political situation in the US. We ask about satisfaction with the Trump administration, direction in which the country is going, and whether respondents would feel safer in Trump's or Biden's America, keying into Trump's narrative that the current unrest is a prelude to a Biden presidency. We furthermore ask a set of questions about threat perception and opinions about debasement of the US presidency. A next block of questions pertains to leadership evaluations, in which both candidates are assessed on an eleven-point Likert scale on 5 personality traits. The campaign evaluation block aims to map news consumption and exposure to ads from or about either candidate, as well as a 11-point grading of both campaigns.

A last module contains COVID-19 related questions, as well as salient issue positions. A Pew Research study showed that the COVID-19 pandemic divided the country more than it previously was (Devlin & Connaughton, 2020). The items related directly to COVID-19 stemmed from Barari's (2020) research about public health messaging at the start of the pandemic in Italy. The salient issue statements were selected by Kieskompas - Election Compass experts from policy platforms, social media positions, and public debates during the campaign. They were taken from Kieskompas - Election Compass's US presidential voting advice application (Etienne, Furstein, et al., 2020).

› **Targeted sampling**

An NBC News/Wall Street Journal poll conducted between September 13-16 2020 found that only 11% of Americans were still undecided (Murray, 2020). In order to analyze ideological patterns and issue positions of a sufficiently large sample of Americans contemplating a Biden vote, we have applied four separate targeted sampling approaches.

At the start of the fielding period, Kieskompas - Election Compass relied on its own non-probability 'opt-in' panel in the US. This panel reached is comprised of 4,760 members who voluntarily signed-up through the fielding of a voting advice application for the 2016 presidential election (Krouwel & Kutiyiski, 2016) and a voting advice application for the 2020 Democratic primary elections (Etienne, Pasquier, et al., 2020). Of those panel members, 685 respondents participated (response rate of 14.4%) and 592 completed the survey (completion rate of 86,4%).

Secondly, Kieskompas - Election Compass advertised its survey on Facebook, targeting specifically the five decisive demographic groups defined in the extant research component of the study.

Thirdly, we made use of Qualtrics' panels. Qualtrics is a leading surveying software, providing access to pre-signed up samples which mirror census representation. For each of the targeted voter groups, they collected 195 responses, with the exception of the Sun Belt voter group, for which 145 responses were collected. A total of 1.051 respondents could be classified as belonging to one of these decisive demographic voter groups, whereas 988 could not.

Voter Group	Facebook criteria	Qualtrics criteria	States	Unweighted n	
White Rust Belt Voters without a College Degree	No college education, moderate & conservative political ideology	White, no college education	PA, MI, WI, OH, IA	207	19.7%
White Sun Belt Voters with a College Degree	College education, moderate & conservative political ideology	White, college education	GA, TX, NC, FL	238	22.6%
Battleground African American voters	Zip codes with >50% African American residents	African American	MI, NC, FL, PA, GA	196	18.6%
Battleground Latino voters	Zip codes with >50% Latino residents	Latino	FL, AZ, NV, TX,	192	18.3%
Battleground youth	18-30 years old	18-30 years old	MI, WI, PA, FL, NC, AZ, OH, GA, TX, NV, IA	218	20.7%
Total				1,051	100%
Not attributed	Not pertaining to any of the above			988	48.5%

Lastly, we relied on snowball sampling, providing each respondent that completed the survey with a link to share among their contacts and a request to do so. In total (including sampling through Facebook and Qualtrics), 1,478 respondents answered the survey through snowballing, Facebook and Qualtrics.

› Weighting

While this study is exploratory and therefore does not aim to make representative claims of the US voting age population, an iterative proportional fitting procedure was applied to the dataset resulting from the survey (unweighted n = 2,039) in order to reduce dominant biases. Through this reduction, the analyses in the upcoming chapter will be more easily interpretable. An iterative proportional fitting procedure was applied to the dataset resulting from the survey (unweighted n = 2,039) on the following parameters:

- › Age (6 categories)
- › race/ethnicity (4 categories: white, African American, Latino, other)
- › sex (2 categories)
- › education (2 categories: low and high education)
- › vote recall for the 2016 presidential elections (5 categories: Trump, Clinton, other, did not vote and was not allowed to vote)

Note that a geographic variable was purposefully not part of the weighting procedure, so as to not negate the focus on battleground states in this study. The population distribution for the demographic parameters was taken from the US Census Bureau (“QuickFacts: United States,” 2010).

In order to assign a weight to a respondent, all of these parameters must be known. This was the case for 1,622 respondents, which is the dataset’s weighted sample size (weight attribution ratio of 79.5%). The weights are trimmed at the 99.9th percentile, resulting in a maximum weight of 11.13 and a design effect of 2.68. The maximum margin of error at the 95% confidence level is 4.0%.

› Selection and categorization of voter groups

In the extant research study, we have identified 5 decisive voter groups, largely based on demographic characteristics and residency in key battleground states. Keeping in mind the aim of this study, from this point onwards, we will expand our analysis to include not only these 5 groups, but to discuss all voters who contemplate a Biden vote.

› Conditional classification

In order to categorize respondents, a number of conditions were drawn up, through which we constructed a 5 category variable defining voter groups by their candidate preference strength, which will underpin most of the analyses performed in the next chapter.

Condition name	Condition components	Unweighted		Weighted	
Core Republican Trump voters	› Identify as a strong Republican › Intention to vote for Trump › Certainty to vote for Trump › Very likely to go vote	220	14.0%	195	13.8%
Considering Trump	› Propensity to vote for Trump 5/10 or higher › Intention to vote for Trump	260	16.5%	353	24.9%
Considering both candidates	› Pro Trump (condition) › Pro Biden (condition)	114	7.2%	132	9.3%
Considering Biden	› Propensity to vote for Biden 5/10 or higher › Intention to vote for Biden	507	32.1%	466	32.9%
Core Democratic Biden voters	› Identify as a strong Democrat › Intention to vote for Biden › Certainty to vote for Biden › Very likely to go vote	476	30.2%	270	19.0%
Total		1577	100%	1415	100%
Not attributed	Not answering to any of the above	462	22.7%	207	12.8%

› Clustering

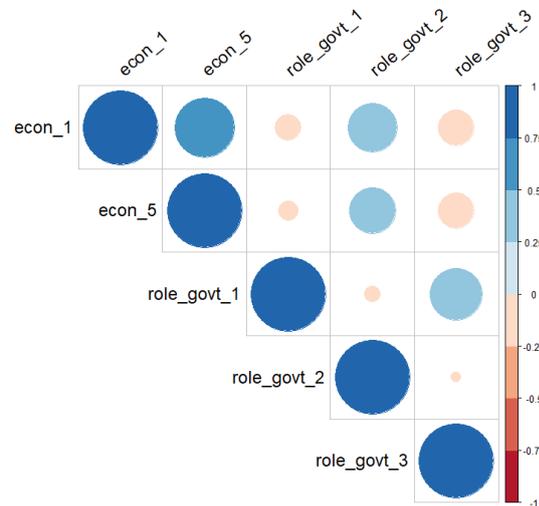
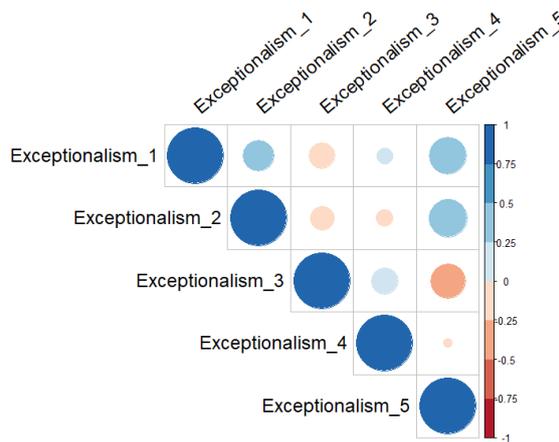
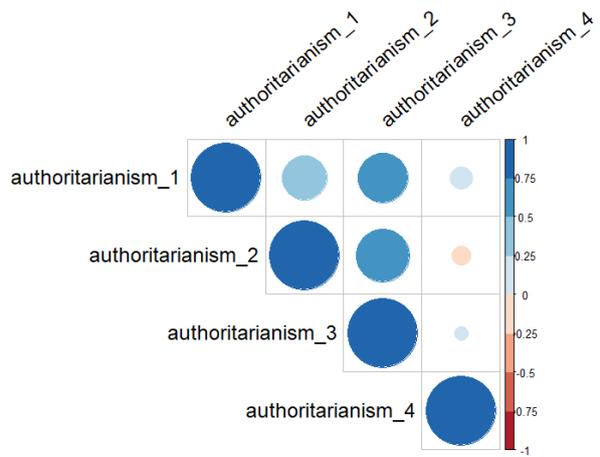
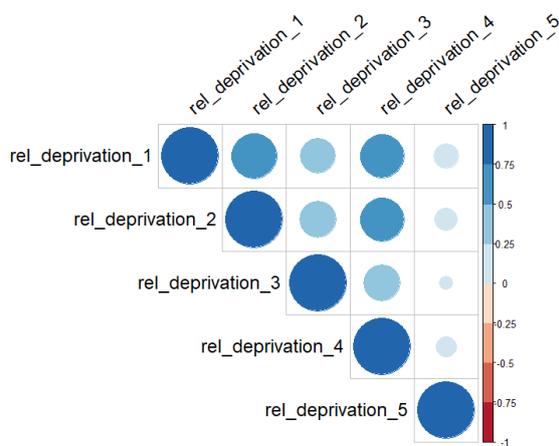
In order to perform more in-depth analysis of voters who consider a Biden vote but are not fully decided, we statistically investigated the voter groups further through cluster analysis. We clustered the combination of respondents from both the group 'Considering both candidates' and the group 'Considering Biden' (n = 621). Although the literature has identified certain groups based on demographics, the cluster analysis was carried out based on psychographics. Psychographics represent an internalized belief or thought pattern, such as values, attitudes, and personality. Psychographic factors lay at the root of our behavior, which is why identifying profiles based on them is increasingly common in marketing, advertisement, and also targeting of political campaigns.

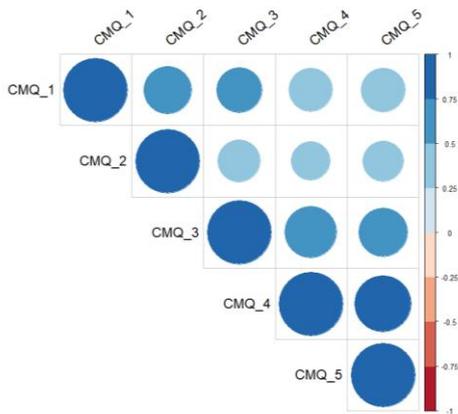
› Cluster analysis dimensions

The psychographic indices used as dimensions in the cluster analysis are authoritarianism, patriotism or exceptionalism, conspiracy beliefs, relative deprivation, and economic ideology. Another question included asks whether the motivation to vote for Biden stems more from a pro-Biden or an anti-Trump view. The reason for selecting these variables instead of opinions on specific topics or a measure of agreement with either candidate is that the analyses suggest that the psychographic measures described above can explain opinion patterns. While clustering on demographics assumes, for example, that all white women between 30 and 50 with a university degree hold similar political beliefs, our analysis assumes nothing of the sort. Rather, it groups people based on similarities in how they value authority, how they perceive America's greatness, how they regard their economic situation, and so on. Thus, individuals similar in worldview are grouped together into clusters within which beliefs and political opinions tend to align.

The indices used as dimensions comprise all items of the respective blocks described in the in the survey subsection of the Methodology. Their statistical reliability and validity were assessed to create robust and unidimensional indices, where all items represent the same underlying factor. Therefore, the index score of an individual is based on the sum of their answers, weighted by the component scores of Principal Components Analysis. It is important to note that for the dimensions used in the cluster analysis, the index-scores were based on PCA of only the cluster subsample of respondents that were considering a Biden vote.

The following correlation plots visualize the relations of items within blocks, or scales. Not all items were retained to create each index. The Relative Deprivation index consists of the first four items in its block. Authoritarianism consists of only the first three items in its block. Exceptionalism consists of four items, since the fourth in the block was left out. The Economic Ideology index consists of the first three items in the block about the role of the government, as well as the first and last item of the economic position statements. Finally, the Conspiracy Belief index contains all original items.





The table below shows how much of the variance of the answers is explained by the first component of the index, and reports Cronbach's alpha as a measure of internal consistency. A value of at least 40% is acceptable for the former; a value higher than 0.60 – 0.70 and lower than 0.95 for the latter. All statistics are based on analysis of the subsample.

Variable	Variance explained by first component	Cronbach's alpha
Relative deprivation	64%	0.81
Authoritarianism	68%	0.76
Exceptionalism	51%	0.68
Economy (new)	43%	0.64
Conspiracy belief	63%	0.85

› Cluster analysis methodology

The cluster analysis is based on the positions of respondents with respect to the aforementioned components. In order to produce a robust segmentation of respondents, we make use of a two-step procedure. The approach to the classification is based on the assumption that, on average, various clustering methods will jointly tend to capture the true clustering structure present in the data. In the first step, a number of candidate clustering structures are generated across a variety of clustering algorithms and the number of clusters. In the second step these results are combined to produce the final classification.

The classification is based on five clustering methods: k-means, hierarchical (two variants: Ward and average linkage agglomeration rules), latent class analysis (across 14 models), fuzzy clustering and k-medoids clustering. Where possible, the clustering is based on both Manhattan and Euclidian distance matrices. The methods are used to generate solutions between 2 and 10 clusters. The best solutions are selected using validation indices. Furthermore, the selected clustering solutions are assessed for similarity by comparison of the produced classifications. Thus, in addition to validation indices, the results are assessed using the degree of overlap and clustering structure. Finally, in order to produce a more robust solution, the most plausible solutions from each algorithm are saved and used to generate a new dissimilarity matrix. The final classification is based on the newly generated distance matrix, where k-medoids clustering is used as a meta-clustering method.

Cluster validation (assessment of optimal number of classes across algorithms) is primarily based on the following set of validation indices: within sum of squares index, average distance within index, average distance between index, separation index, average within/average between ratio. Furthermore, where possible following indices are used: gap statistics, KL Krzanowski and Lai (1988) index, Calinski and Harabasz index, Hartigan index, cubic clustering Criterion index, Scott index, Marriot index, trace of within clusters pooled, covariance matrix index, Tracew index, Friedman index, Rubin index, Cindex, Davies and Bouldin DB (1979) index, silhouette index, Duda index, Pseudot2 index, Beale index, Ratkowsky index, Ball index, point-biserial index, Frey index, McClain index, Dunn index, Hubert index, SDindex, Dindex, SDbw index, and Bayesian information criterion. The indices are calculated on the Euclidian dissimilarity matrix. For each clustering method the candidate results are selected using majority vote.

› Final clustering solution

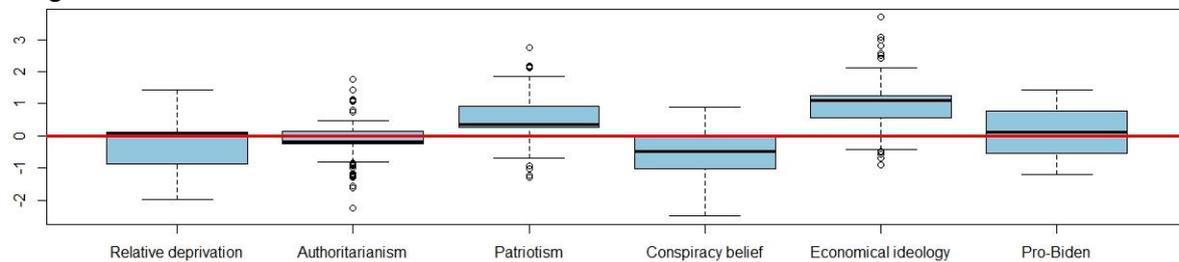
The validation indices indicate that most clustering methods produce the best solutions by dividing respondents into 4 classes. Significant exceptions are fuzzy clustering and hierarchical average linkage agglomeration clustering, which suggest divisions into a large number of clusters, namely 9 and 10 classes. Upon the comparison and review of the results, we have selected four distinct clustering solutions which separate respondents into four classes for further analysis: k-medoids based on Manhattan dissimilarity matrix; k-means; and latent class solutions based on EII and VII constraints. The similarity of the selected solutions ranges between approximately 72.4% and 93.6% (i.e. the respective share of cases is grouped in the corresponding classes of the alternative classification). These solutions are combined via distance matrix which was used in k-medoid clustering to obtain the final classification. The final solution classifies between 82% and 93% cases in the clusters which correspond to the classes produced by algorithms used to generate the dissimilarity matrix.

› Cluster segmentation and interpretation

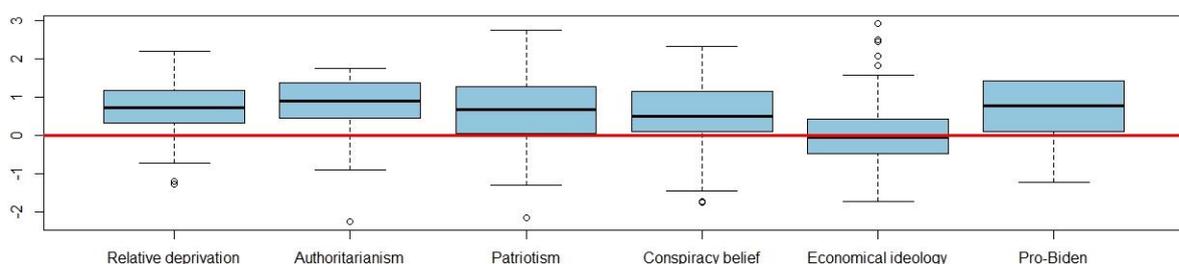
This final solution to the cluster analysis divides the subsample in four distinct clusters. Boxplots representing the scores within each cluster on the included psychographic dimensions can be found underneath. Higher dimension scores are associated with higher relative deprivation, higher authoritarianism, higher patriotism, stronger belief in conspiracies, and more economically right-wing positions. The Pro-Biden dimension measures whether a respondent's Biden vote would be more pro-Biden or more anti-Trump, with higher scores signifying a more pro-Biden measurement. The red line indicates the unweighted average score of the subsample on the dimensions; as the scores of each dimension are the principal component scores for an index, the average is zero.

The first cluster will further on be referred to as the Unaligned Doubters, the second cluster will be referred to as the Anti-Trump Progressives, the third as the Struggling Moderates, and the fourth as the Intellectual Liberal Moderates. This interpretation will be further substantiated by in the description of the analyzed groups subchapter.

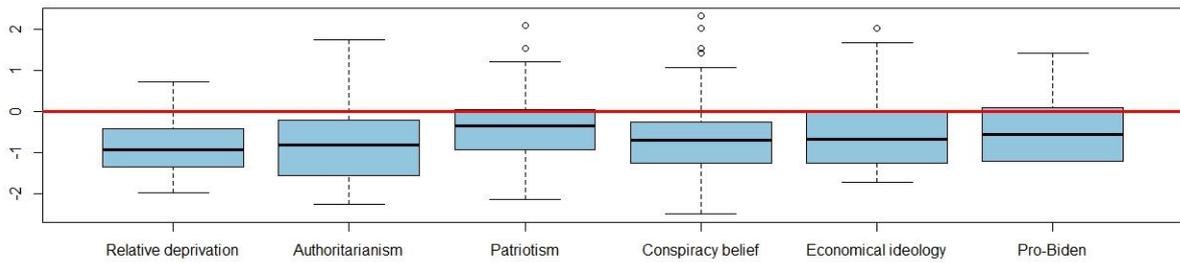
› Unaligned Doubters



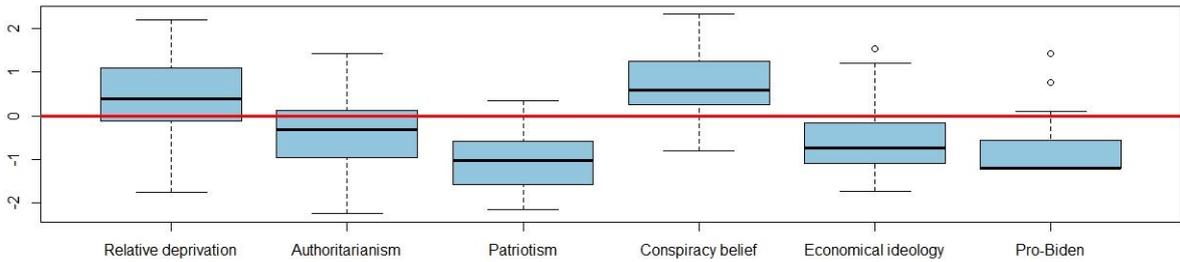
› Struggling Moderates



› **Intellectual Liberal Moderates**

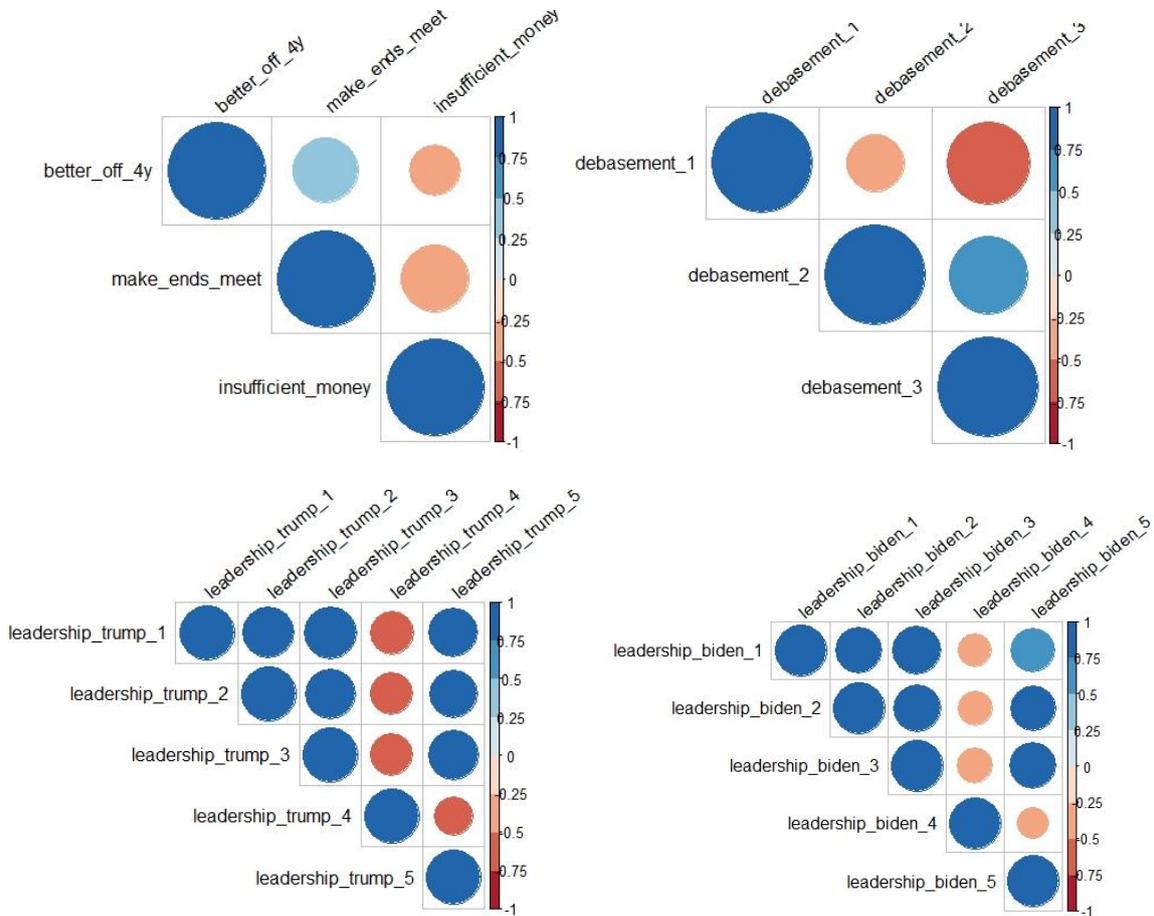


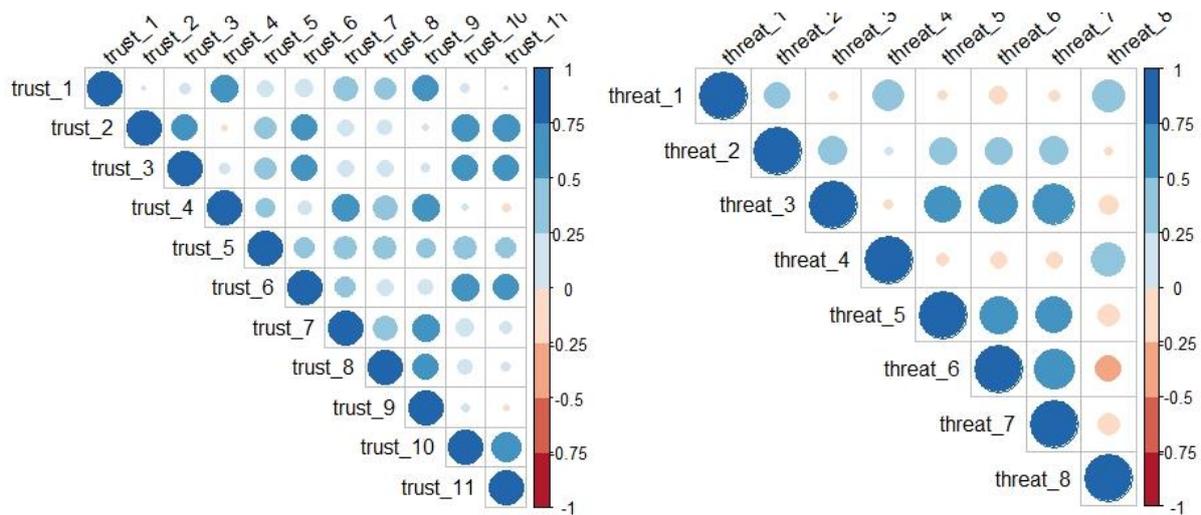
› **Anti-Trump Progressives**



› **Analysis**

The indices that are used in further analysis are created in a similar manner as the indices that served as dimensions for the cluster analysis. A major difference is that the principal components scores that indicate the total score on an index, are based on PCA of the entire sample instead of the subsample. Correlation plots illustrate the correlations between items in blocks.





The Precarity index consists of three items, representing the respondents' current financial situation, compared to the past, and a prediction for the next year. Both Leadership indices contain all characteristics both candidates were rated on. The items in the Trust block were divided in two separate groups. The first index contains items measuring trust towards more progressive entities (e.g., NATO, the media), while the second on more conservative entities (e.g., the military; the justice system). The items in the Threat block also were separated to form two new indices: threat perception of more conservative entities (e.g., Russia, white supremacy) and threat perception of more progressive entities (e.g., Antifa, BLM, socialism). Finally, the Debasement index contains all original three items in its block.

The following table shows the validity of the indices, again based on how much of the variance is explained by the first component and the Cronbach alpha of the index. These analyses are based on the entire sample.

Variable	Variance explained by first component	Cronbach's alpha
Precarity	61%	0.67
Leadership Trump	80%	0.94
Leadership Biden	75%	0.91
Trust (conservative)	56%	0.84
Trust (progressive)	65%	0.87
Threat (conservative)	62%	0.85
Threat (progressive)	63%	0.71
Debasement	70%	0.78

› Bibliography

- ANES. (2017). *2016 Time Series Study - ANES | American National Election Studies*.
<https://electionstudies.org/data-center/2016-time-series-study/>
- Antonucci, L., Horvath, L., Kutiyiski, Y., & Krouwel, A. (2017). The malaise of the squeezed middle: Challenging the narrative of the 'left behind' Brexiter. *Competition & Change*, 21(3), 211–229.
<https://doi.org/10.1177/1024529417704135>
- Barari, S., Caria, S., Davola, A., Ivchenko, A., Jachimowicz, J., King, G., Kraft-todd, G., Ledda, A., MacLennan, M., Falco, P., Fetzer, T., Fiorin, S., Hensel, L., Ivchenko, A., Jachimowicz, J., King, G., Kraft-todd, G., Ledda, A., MacLennan, M., ... Slepoy, F. R. (2020). *Evaluating COVID-19 Public Health Messaging in Italy: Self-Reported Compliance and Growing Mental Health Concerns*. 1–19. .
- Bruder, M., Haffke, P., Neave, N., Nouripanah, N., & Imhoff, R. (2013). Measuring Individual Differences in Generic Beliefs in Conspiracy Theories Across Cultures: Conspiracy Mentality Questionnaire. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00225>
- Curtice, J., Hudson, N., & Montagu, I. (2020). *British Social Attitudes: The 37th Report*.
- Davidov, E. (2009). Measurement equivalence of nationalism and constructive patriotism in the ISSP: 34 countries in a comparative perspective. *Political Analysis*, 17(1), 64–82.
<https://doi.org/10.1093/pan/mpn014>
- Devine, P. (2007). *Northern Ireland Life and Times Survey 2007*. <https://www.ark.ac.uk/nilt/2007/tech07.pdf>
- Devlin, K., & Connaughton, A. (2020). *Most Approve of National Response to COVID-19 in 14 Advanced Economies*. <https://www.pewresearch.org/global/2020/08/27/most-approve-of-national-response-to-covid-19-in-14-advanced-economies/>
- Elchardus, M., & Spruyt, B. (2014). Populism, Persistent Republicanism and Declinism: An Empirical Analysis of Populism as a Thin Ideology. In *Government and Opposition* (Vol. 51, Issue 1, pp. 111–133). Cambridge University Press. <https://doi.org/10.1017/gov.2014.27>
- Etienne, T., Furstein, R., & Krouwel, A. (2020). *2020 US Presidential Election Compass*. Kieskompas - Election Compass. presidential.electioncompass.org
- Etienne, T., Pasquier, A., Furstein, R., & Krouwel, A. (2020). *US Democratic Primaries 2020 Election Compass*. Kieskompas - Election Compass. usa.electioncompass.org
- Hameleers, M. (2020). Blaming in the name of our people: how attitudinal congruence conditions the effects of populist messages communicated by traditional media, politicians, and citizens. *Media Psychology*, 1–22. <https://doi.org/10.1080/15213269.2020.1785314>
- Krouwel, A., & Kutiyiski, Y. (2016). *2016 US Presidential Election Compass*. Kieskompas - Election Compass.
- Murray, M. (2020, September 20). After a tumultuous month of news, Biden maintains national lead over Trump. *NBC News*. <https://www.nbcnews.com/politics/meet-the-press/after-tumultuous-month-news-biden-maintains-national-lead-over-trump-n1240501>
- QuickFacts: United States. (2010). In *U.S. Census Bureau*.
<https://www.census.gov/quickfacts/fact/table/US/PST045219>
- Smith, T. W. (1992). The international social survey program. *International Journal of Public Opinion Research*, 4(3), 1992.
- Walker, I., & Smith, H. J. (2002). *Relative Deprivation Specification, Development, and Integration*.
<http://www.cambridge.org>