Appendix

Neither a trait nor wildly fluctuating: on the stability of populist attitudes and its implications for empirical research

Christian H. Schimpf[†], Alexander Wuttke[‡], Harald Schoen[‡‡]

[†]University of Saskatchewan; [‡]LMU Munich; [‡‡]University of Mannheim

Corresponding email address: *christian.schimpf@usask.ca*

Contents

Appendix 1 - Information about the panel, question wording and descriptive statistics

This section contains the question wording for variables other than populist attitudes (see main text) along with descriptive statistics for all variables.

Additional Panel Information

The GLES Panel Study (GLES 2021) consisted of multiple sample components, including refresher samples added as time went on. Four our analyses, we considered three sample components namely, A1 (original 2017 election campaign sample), A2 (respondents who had participated in the 2013 GLES election campaign panel) and A3 (refresher sample component added between GLES Panel Waves 3 and 4).

In the paper, we use running numbers to label the waves from one through six. The numbers correspond to the original panel wave number as seen in Table A1 (by variable).

Table A1.1 Correspondence of wave numbering in the paper to original panel waves in the GLES dataset

Running Number in paper Variable	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Populist attitudes	Wave 5	Wave 8	Wave 9	Wave 13	Wave 14	Wave 15
Political interest	Wave 5	Wave 8	Wave 9	Wave 13	Wave 14	Wave 15
Satisfaction with government performance	Wave 5	Wave 8	Wave 13	Wave 14	Wave 15	
Economic perceptions	Wave 5	Wave 8	Wave 13	Wave 14	Wave 15	
Climate change protection	Wave 4[1]	Wave 8	Wave 13	Wave 14	Wave 15	

---

[1] For the refresher sample component, entries were taken from Wave a1.

Question Wording

The questionnaire was devised in German. English translations are taken from the official translation provided by the German Longitudinal Election Project. In the following, "X" is a placeholder for the wave number in the variable name. For all scales, opinion boundaries were determined by the scales' midpoints such that, for example, we counted moving from "interested" (very/somewhat) to "in between" on the political interest variable as having crossed an opinion boundary.

Political Interest (Variable name: kpX_010): Quite generally, how interested are you in politics?

Coding
1. Very interested
2. Somewhat interested
3. In between
4. Not very interested
5. Not at all interested

Satisfaction with government performance (kpX_730): The wording was identical for all but one wave (originally Panel Wave 8, here labeled as Wave 2).

V1 - Are you more satisfied or less satisfied with the performance of the government of CDU/CSU and SPD?

V2 (Wave 2) - Are you more satisfied or less satisfied with the performance of the government of CDU/CSU and SPD in the last four years?

Original Coding	Coding in this paper
-5 Not at all satisfied	1 Completely satisfied
…	…
+5 Completely satisfied	5 Not at all satisfied

Note: For convenience purposes we recoded and collapsed the original categories as follows: -5 to -3=5, -2 and 1- =4, 0=3, 1 and 2=2, 3 to 5=1.

Economic Perceptions (kpX_820): How would you evaluate the current general economic situation in Germany?

Coding
1. Very good
2. Good
3. Neither good nor bad
4. Bad
5. Very bad

Climate protection (kpX_1290): Part of the question wording (the question part at the end) varied across different panel waves; The original questionnaire only included value labels for the two extreme points

Some say that the fight against climate change should definitely take precedence, even if it impairs economic growth. Others say that the economic growth should definitely take precedence, even if it impairs the fight against climate change.

V1-What is your personal view on climate change and economic growth? (Original panel waves 4, and 13-15; corresponding here to 1, 3-5)

V2-What is your personal view on this issue? (Original panel wave 8, here 3)[2]

Coding
1 fight against climate change should take precedence, even if it impairs economic growth
…
7 economic growth should take precedence, even if it impairs the fight against climate change

Descriptive Statistics

Table A1 shows the descriptive statistics and also includes information on the number of valid responses by panel wave and the per cent of missing cases, calculated with the full dataset of N=22,543 as basis. Note that for populist attitudes, higher values imply *less* agreement with populist ideas.

---

[2] The English Questionnaire provided by GLESindicates more than two different question wordings. However, the German questionnaire only indicates two different wordings which is what we base our information here.

Table A1.2 Descriptive Statistics and Missing Data

Variable	Statistics	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
	Valid N	13,542	12,848	11,672	9,222	8,312	8,323
Populist attitudes	% missing	39.93%	43.00%	48.22%	59.1%	63.13%	63.08%
	Mean	2.87	2.83	2.85	2.81	2.79	2.76
	Std. Dev.	0.75	0.73	0.77	0.74	0.75	0.77
	N of respondents with valid answers in all six waves: 5,333						
	Valid N	14,125	13,372	12,224	9,552	8,653	8,621
Political interest	% missing	37.34%	40.68%	45.77%	57.63%	61.61%	61.76
	Mean	2.58	2.48	2.5	2.5	2.5	2.52
	Std. Dev.	1.06	1.04	1.03	1.05	1.07	1.05
	N of respondents with valid answers in all six waves: 6,231						
	Valid N	14,060	13,260	9,505	8,576	8,556	
Satisfaction with government performance	% missing	37.63%	41.18%	57.84%	61.96%	62.05%	
	Mean	3.21	3.25	2.78	2.88	3.16	
	Std. Dev.	1.5	1.5	1.52	1.54	1.52	
	N of respondents with valid answers in all six waves: 6,426						
	Valid N	13,958	13,237	9,470	8,538	8,532	
Economic perceptions	% missing	38.08%	41.28%	57.99%	61.13%	62.15%	
	Mean	2.5	2.39	3.19	3.23	3.3	
	Std. Dev.	0.86	0.81	0.9	0.83	0.85	
	N of respondents with valid answers in all six waves: 6,381						
	Valid N	14,929	13,180	9,432	8,494	8,473	
Climate change protection	% missing	33.78%	41.53%	58.16%	62.32%	62.41%	
	Mean	3.1	3.23	3.49	3.53	3.54	
	Std. Dev.	1.55	1.55	1.58	1.59	1.6	
	N of respondents with valid answers in all six waves: 6,324						

Measurement Error at the Dimension Level

For our analyses, we use the mean scores across the three items for each of the dimensions. To ensure that this simple calculation does not introduce measurement error that might affect our results, we also ran confirmatory factor analyses for each of the dimensions at each wave and calculated the correlation between means and predicted scores based on the confirmatory factor analyses. The table below shows that there are almost not differences between the means and predicted CFA scores. Freeder et al. (2019) find for correlations of policy opinions over time stabilize the more items one has available. Applied to this context, the finding suggests that three

items sufficiently reduce measurement error such that proceeding with the mean scores does not introduce imprecision to a level that would affect our analyses in a substantial way.

Table A1.3 Correlations between means across three items and predicted Confirmatory Factor Analyses scores for each of the dimensions at each panel wave

Dimension Panel Wave	Anti-Elitism	People Sovereignty	Homogeneity
1	0.943	0.991	0.979
2	0.95	0.99	0.976
3	0.955	0.992	0.976
4	0.954	0.993	0.982
5	0.969	0.993	0.98
6	0.958	0.995	0.977

Appendix 2 - Measurement Invariance

We check whether the measures for our three dimensions, anti-elitism, homogeneity, and sovereignty are invariant over time. We do so by fitting four CFA models, each of which adds an additional constraint to the model. In the first model, the configural model, we allow factor loadings, item intercepts, and residual variances to vary across panel waves. For the second model, the metric model, we fix factor loadings to be equal across panel waves. In the third model, the scalar model, we further force intercepts to be equal over time. Scalar invariance is particularly important in our context because it ensures that absolute levels have the same meaning over time (Mackinnon and Curtis 2020, 5). Lastly, we fit a model to test for residual invariance by fixing residual item variation to be equal across panel wavs.[3] To compare the models, use the Comparative Fit Index (CFI) value and the root mean square error of approximation (RMSEA). Both have been found to work reasonably well in assessing invariance (Chen 2007; Cheung and Rensvold 2002).[4] We also report other fit measures to evaluate the overall model fit (Table A2.1) The results show that all types of models fit the data reasonably well. CFI values above 0.95 (larger values reflect better model fit) and RMSEA values below 0.05 (smaller values indicate better model fit) are indicators of good fit, all of which we see here (Hu and Bentler 1999; Kline 2005). The next step is to assess fit across models. Any added constraint usually results in worse fit. The crucial question is whether fit deteriorates substantially enough to prefer the more parsimonious model over the model with added constraints. As general markers, the more constrained model is preferred to the more parsimonious model if CFI values decrease by 0.01 of less (Cheung and Rensvold 2002) and RMSEA values increases by no more than .015 (Chen 2007). We find that these cut-off points hold across all three dimensions and model configurations, leading us to conclude that our measures for the three dimensions are invariant over time.

---

[3] For detailed discussions, see Mackinnon and Curtis 2020 or Liu et al. 2017.
[4] Notably absent here is the popular chi-square comparison because it has been shown to be overly sensitive when using larger samples, like the one we work with, often resulting in Type I errors (Cheung and Rensvold 2002).

Table A2.1 Measurement Invariance Test for Populist Attitudes (Listwise deletion of missings)

Dimension		Configural	Metric	Model Scalar	Residual
Anti-Elitism	No. of estimated parameters	109	99	89	74
	Raw loglikelihood	-112788.774	-112808	-112879	-112950
	Robust CFI	0.997	0.996	0.994	0.992
	Δ Robust CFI	NA	-0.001	-0.002	-0.002
	Robust RMSEA	0.02	0.02	0.025	0.027
	Δ Robust RMSEA	NA	0	0.005	0.002
Homogeneity	No. of estimated parameters	109	99	89	74
	Raw loglikelihood	-108494.521	-108500	-108627	-108730
	Robust CFI	0.995	0.995	0.992	0.989
	Δ Robust CFI	NA	0	-0.003	-0.003
	Robust RMSEA	0.026	0.025	0.031	0.034
	Δ Robust RMSEA	NA	-0.001	0.006	0.003
Sovereignty	No. of estimated parameters	109	99	89	74
	Raw loglikelihood	-106715.446	-106735	-106801	-106870
	Robust CFI	0.994	0.993	0.992	0.991
	Δ Robust CFI	NA	-0.001	-0.001	-0.001
	Robust RMSEA	0.035	0.034	0.035	0.036
	Δ Robust RMSEA	NA	-0.001	0.001	0.001

Note that we treat each of the latent dimensions as continuous even though we use a five-point likert scale. Our choice for treating the latent dimensions as continuous is that using five or more categories usually ensures that models are unbiased, a problematic that arises with four or less categories (c.f. Liu et al. 2017, 488).

Appendix 3 - Test to identify whether changes in populist attitudes are purely due to measurement error.

One question that emerges is whether our findings of the two groups (stable and unstable populist attitude holders) is the result of measurement error which might obfuscate that populist attitudes are actually stable for most respondents at the individual level. In addressing this issue, we apply a test akin to what Hill and Kriesi (2001a, 2001b) proposed, following the classification of respondents by means of a finite mixture model (c.f. Kiley 2021). If the observed groups were the consequence of measurement error, we would expect that the average within-person standard deviations across the groups are virtually identical. Table A3.1 shows that the average within-person standard deviations are smaller among respondents classified here as stable compared to the group of unstable respondents, lending additional support to the idea that there is, at the very least, no single group of respondents which holds stable populist attitudes.

Table A3.1 Comparing the average within-person standard deviation of populist attitudes between respondents classified as stable populist attitude holders and respondents classified as unstable populist holders.

Group	Average Within-Person Standard Deviation
Unstable	0.46
Stable	0.28

Appendix 4 - Considering the middle-category as indicating ambivalence/lack of opinion

In the main analyses, we assumed that respondents who neither agree nor disagreed with populist ideas (aggregated populism score of equal to or greater three and smaller four) have a nuanced opinion on populist ideas. As a consequence, we treated the aggregated populism measure as a quasi-continous variable. However, like the concepts we chose for comparison, the original question items for each populist dimension were measured on five-point likert scales. In the literature on likert scales and stability, some argue and demonstrate that people hovering around the midpoint of likert scales are either ambivalent or lack a strong opinion, that is, they do not hold a nuanced opinion (e.g. Kiley 2021, Freeder 2019). To test how sensitive our results are to changing the underlying assumption,. we replicate some of our primary analyses. Note that we do not replicate the correlation to assess relative stability. In principle, our correlation measure, Pearson's r, assumes a metric variable, which would be more suited for our treatment of the middle category as an increment. Yet, the middle category barely affect our correlations because of how its calculated (c.f. Freeder et al 2019, Appendix) so we focus on our tests of absolute stability and subsequent analyses of how stability might affect how we study populist attitudes.

First, we re re-run our main analysis, categorizing those who consistently fall into the "neither agree nor disagree" category as unstable (as opposed to stable under the viewpoint that these are nuanced attitudes). The overall result of more respondents being less stable than stable for all comparative constructs but economic evaluations still hold (Figure A4.1) compared to the findings in the main text (Figure 3 in the main text).[5] However, relative to all other concepts but political interest, populist attitudes become less stable. The reason is that the per cent of respondents holding stable neither agree nor disagree attitudes is the second highest for populist attitudes at 11%. By comparison, only 2% (Satisfaction with government), 6% (Economic perceptions), and 5% (Climate change) fall consistently into the middle category for all concepts but political interest (the per cent of respondents in the middle category is 12% here). Thus, while the overall conclusion that populist attitudes is neither fully stable nor wildy fluctuating is still

---

[5] Note that we not only changed the assumption of how to categorize neither agree nor disagree responses, that is the middle category, for populist attitudes but also for all other selected concepts here to ensure comparability.

valid, considering the middle category as ambivlance or lack of opinion shows some effect with downstream consequences for analyses as we elaborate further below.

Figure A4.1 Stability of populist attitudes over time in comparison to selected concepts under assumption that respondents who neither agree nor disagree with populist attitudes are ambivalent/lack a clear opinion
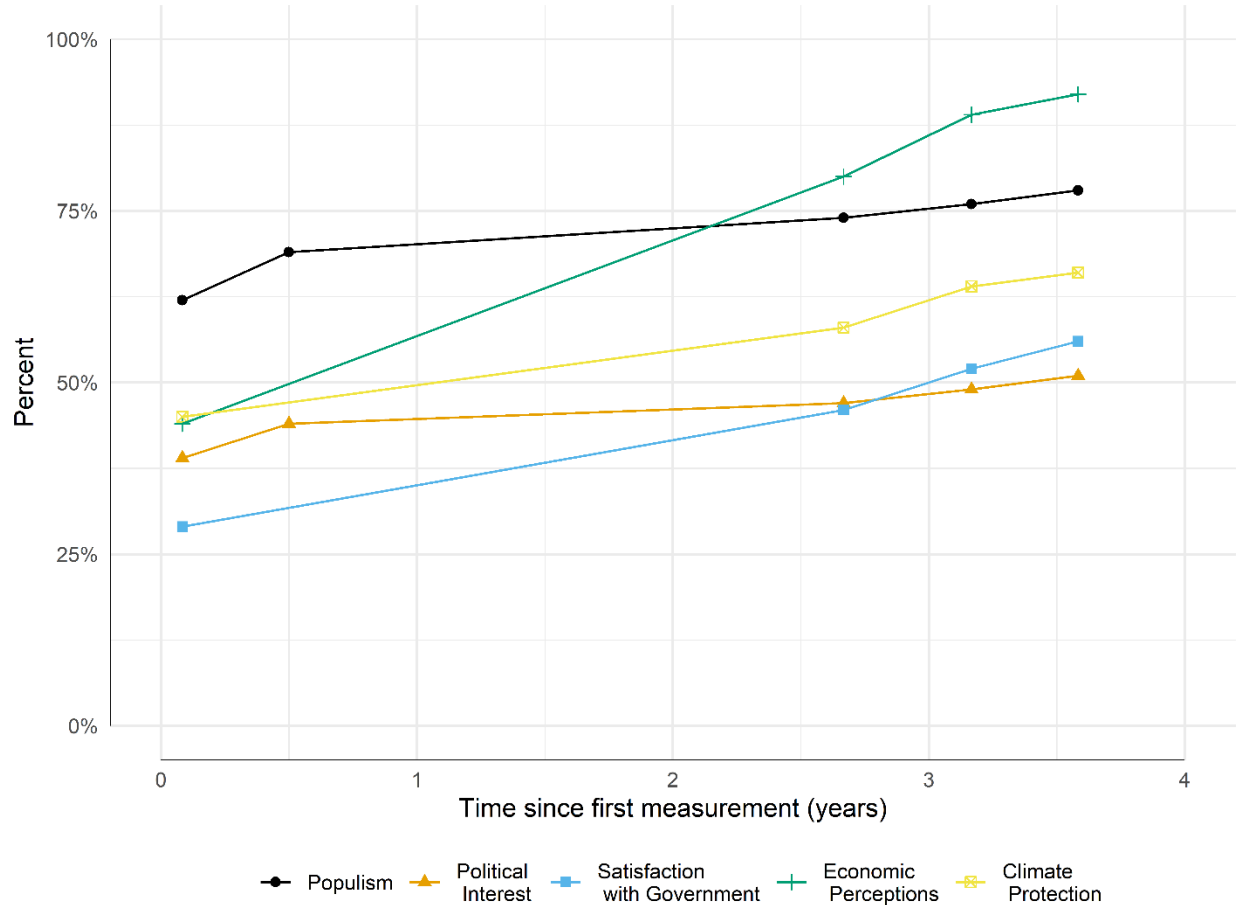

Second, we look at how often respondents cross opinion boundaries in Table A4.1 using the alternative categorization of the middle category (akin to Table 2 in the main text). The results show two things. First, the number of respondents who are stable in their opinions decreases by 11.58% which is the share of respondents who consistently neither agree nor disagree. This share of respondents is roughly equivalent in seize to the respondents Kiley (2021, 12) identifies as consistently neither agreeing nor disagreeing with gay marriage in the US based on GSS data from 2006-2014. Second, the number of of times respondents cross opinion boundaries of one or more times is more concentrated at the higher end because neither agreeing nor disagreeing in two consecutive panel waves or more is considered as unstable.

Table A4.1 Share of respondents by number of times they cross opinion boundaries under assumption that respondents who neither agree nor disagree with populist attitudes are ambivalent or lack an opinion

Number of times a respondent crosses an opinion boundary	0	1	2	3	4	5
Percent of respondents	22.22	3.96	10.56	11.63	14.46	37.18

Third, we also replicated our correlational analyses with attitudes towards democracy and self-reported vote choice. Figure A4.2 shows the correlation between the three items capturing attitudes towards democracy and populist attitudes by stability (akin to Figure 4 in the main text; identical set of respondents included in the analyses). In our main analyses, we saw that the correlations differed from stable opinion holders from those with unstable attitudes towards populism and also, the overall set of respondents. We see a similar pattern here with differences in correlations between the two groups and the combined set of respondents.

Figure A4.2 Correlations between populist attitudes and items on liberal democracy by stability (considering stable respondents in the middle category as holding a nuanced opinion.


12
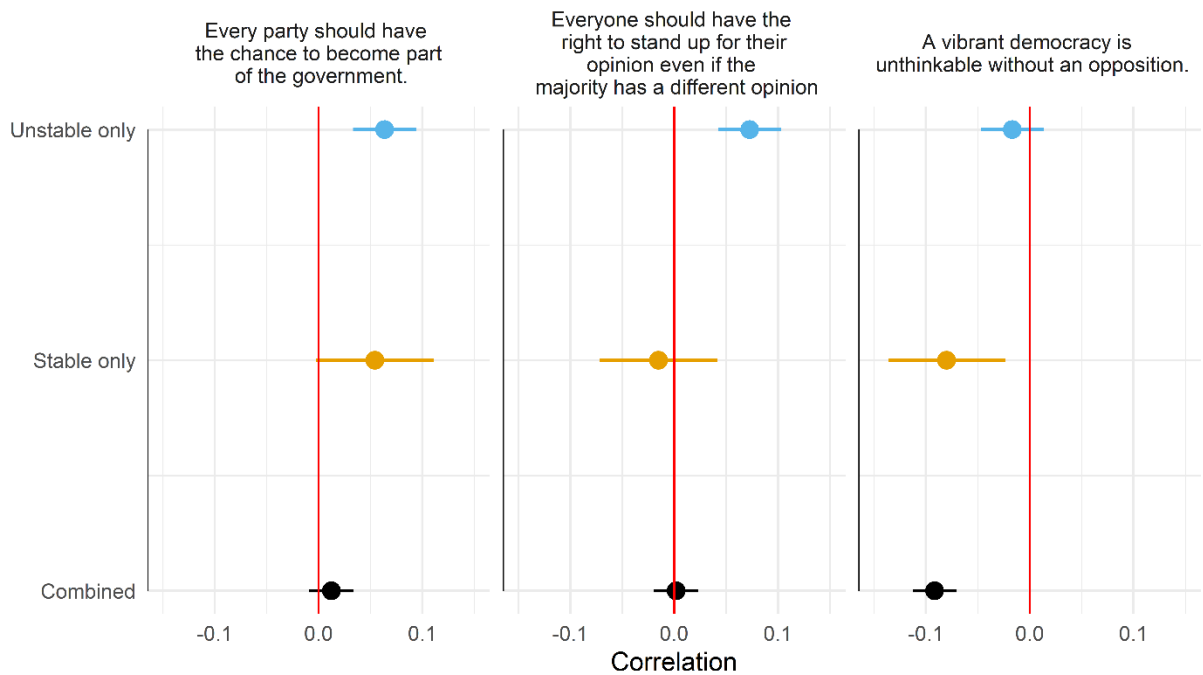
Appendix 5 – Assessing stability across the dimensions of populist attitudes and sensitivity to different aggregating mechanisms

Stability of individual populism dimensions

Figure A6.1 shows the correlations over time between the measures of the individual dimensions at the first time of measurement and subsequent panel waves. The graph shows that *People's Sovereignty* is the most stable of the three dimensions and *Homogenous People* is the least stable in terms of relative stability.

Figure A5.1 Correlations between measures taken at initial wave and subsequent waves for individual dimensions of populist attitudes.
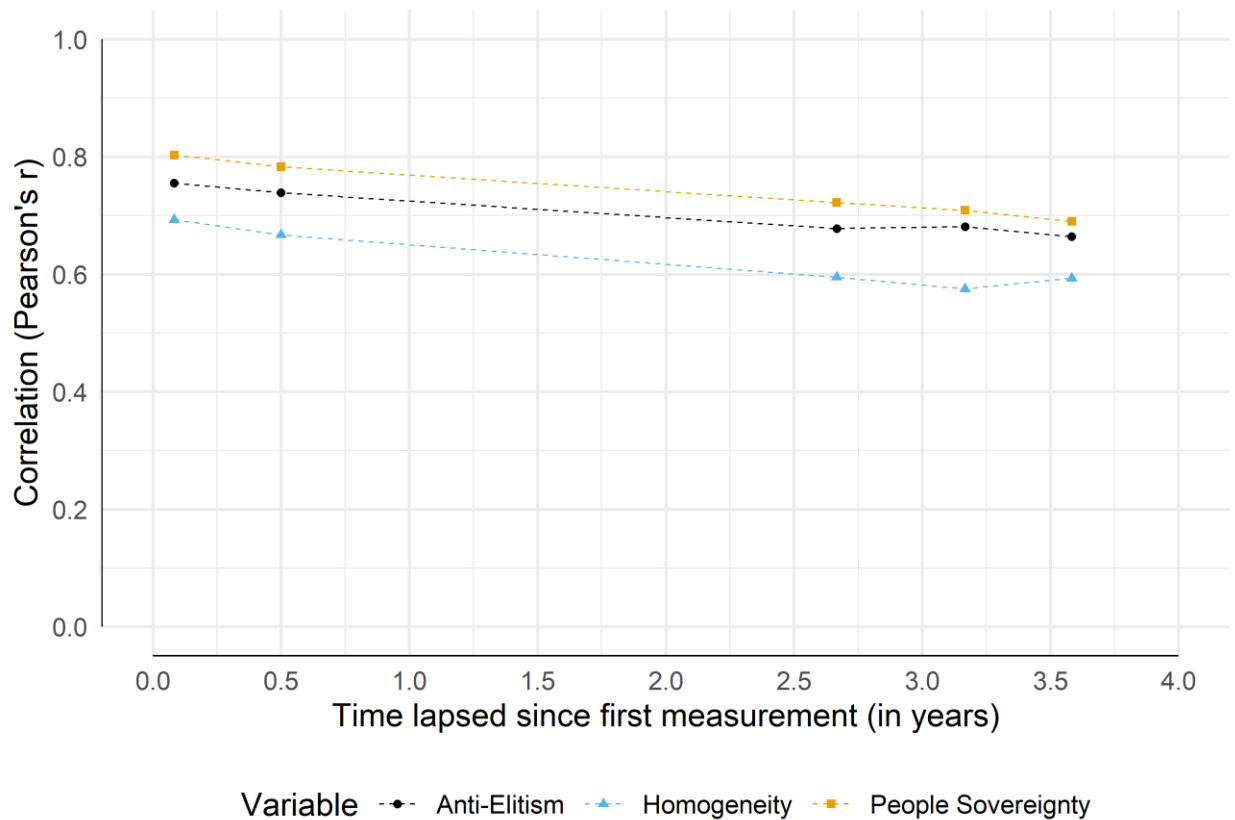
Looking at absolute stability (Figure A5.2), we find the homogeneity dimension has the largest number of respondents with unstable attitudes towards the dimension's ideas, followed by the sovereignty dimension and the anti-elitism dimension. Comparing the individual dimensions to our aggregate measure in the main text, we see that the results for the aggregate score most closely resembles the homogeneity dimension, both with respect to relative and absolute stability. The

reason is that the homogeinity score of individual respondents is most frequently the one panelists agree with the least, resulting in the minimum score to originate from the homogeneity dimension.[6]

Figure A5.2 Share of respondents with unstable attitudes towards individual dimensions of populism over time
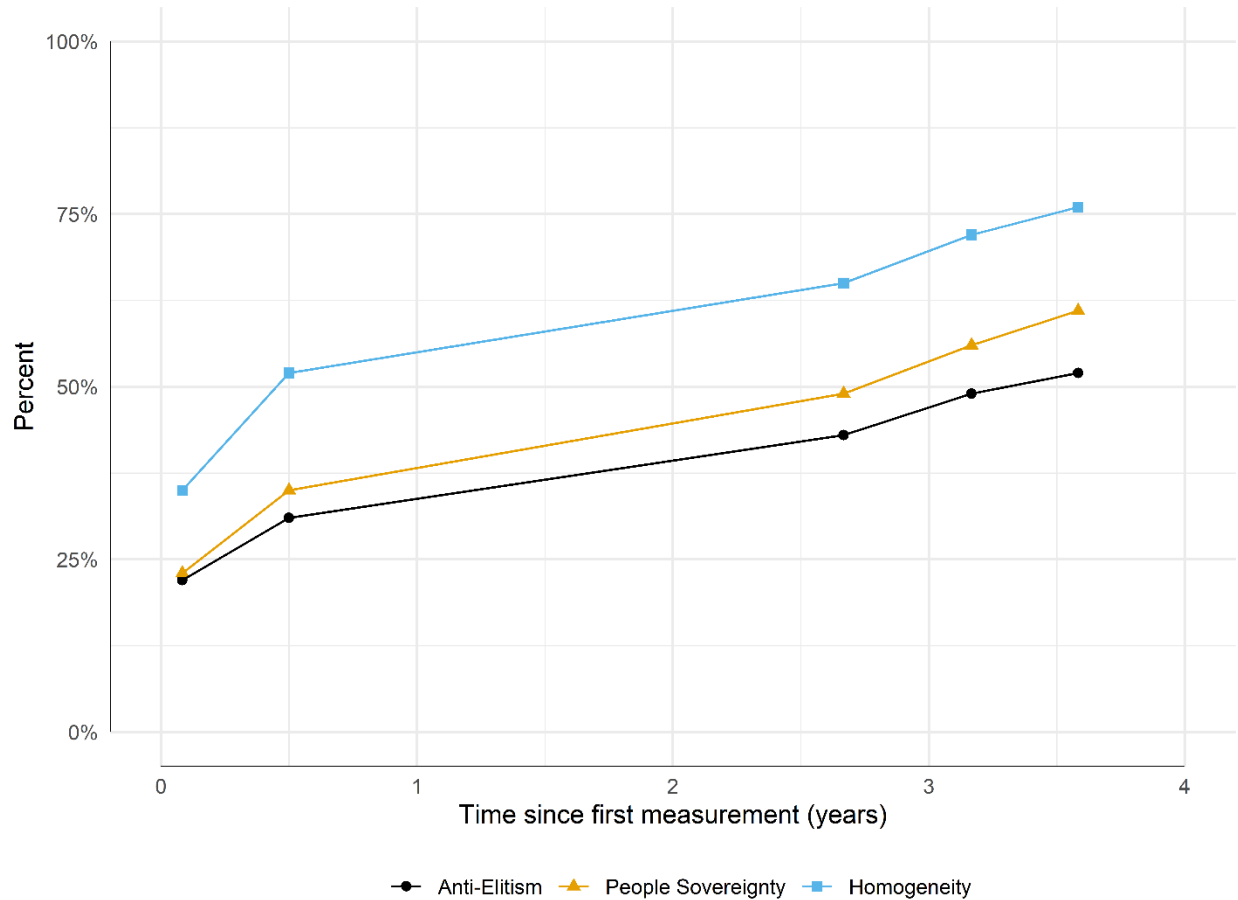
Similarly, we see that when using absolute changes as a measure of absolute stability, anti-elitism, on average, shows the least movement whereas homogeneity shows the most instability (Table 5.1).

---

[6] For all respondents included in the main analyses (N=5,333), the homogeneity score is the single lowest score on the three dimensions in 50% to 62.7% across the six panel waves. In addition, the homogeneity score is as low as the score on another dimension in 16.6% to 18.4% of the panel waves.

Table A5.1 Absolute changes in attitudes and average absolute within-respondent change across subsequent panel waves for individual populist dimensions (standard deviation in parentheses)

	Mean Absolute Change (sd)					
Populist dimensions	W1-W2	W2-W3	W3-W4	W4-W5	W5-W6	Across pairs of panel waves
Anti-Elitism	0.37	0.38	0.48	0.37	0.36	0.39
	(038)	(0.38)	(0.44)	(0.38)	(0.38)	(0.23)
Homogeneity	0.45	0.46	0.51	0.45	0.43	0.46
	(0.44)	(0.45)	(0.49)	(0.45)	(0.43)	(0.26)
Sovereignty	0.39	0.39	0.49	0.39	0.38	0.41
	(0.42)	(0.42)	(0.49)	(0.41)	(0.4)	(0.26)

Using multiplication to aggregate the populism dimension scores

Multiplying the individual dimension scores to obtain an aggregated score of populism still adheres to the principle of non-compensation though to a lesser degree than using the minimal score. Thus, high scores on one dimension can somewhat offset lower scores on other dimensions which is not possible when using the minimal score. Using multiplication thus reduces the influence of the homogeneity score on our aggregate measure. This can aligne the results from stability analyses closer to the people's sovereignty and anti-elitism dimension. For relative stability (Figure 5.3) and absolute stability (Figure 5.4), we see that the multiplied score suggests a more stable construct of populism compared to the minimal score.[7]

---

[7] As in the main text, we recoded respondents into three categories based on the continuous multiplied populism score (Range: 1 to 125; lower scores reflected higher agreement with populist ideas in our coding). To determine the agreement, we used equal cut-off points to determine agreement, disagreement, or ambivalent opinions. For example, respondents who scored below 15.625 (akin to using 2.5 on a five-point likert scale as 2.5 to the power of three=15.265) were coded as agreeing with the ideas of populism. Respondents who scored equal to or above 42.825 were coded as disagreeing with the ideas of populism. We then determine whether respondents continuously agreed, disagree, or changed their attitudes across the panel waves.

Figure A5.3 Correlations between measures taken at initial wave and subsequent waves for populist attitudes (minimum score and multiplied) and individual dimensions.
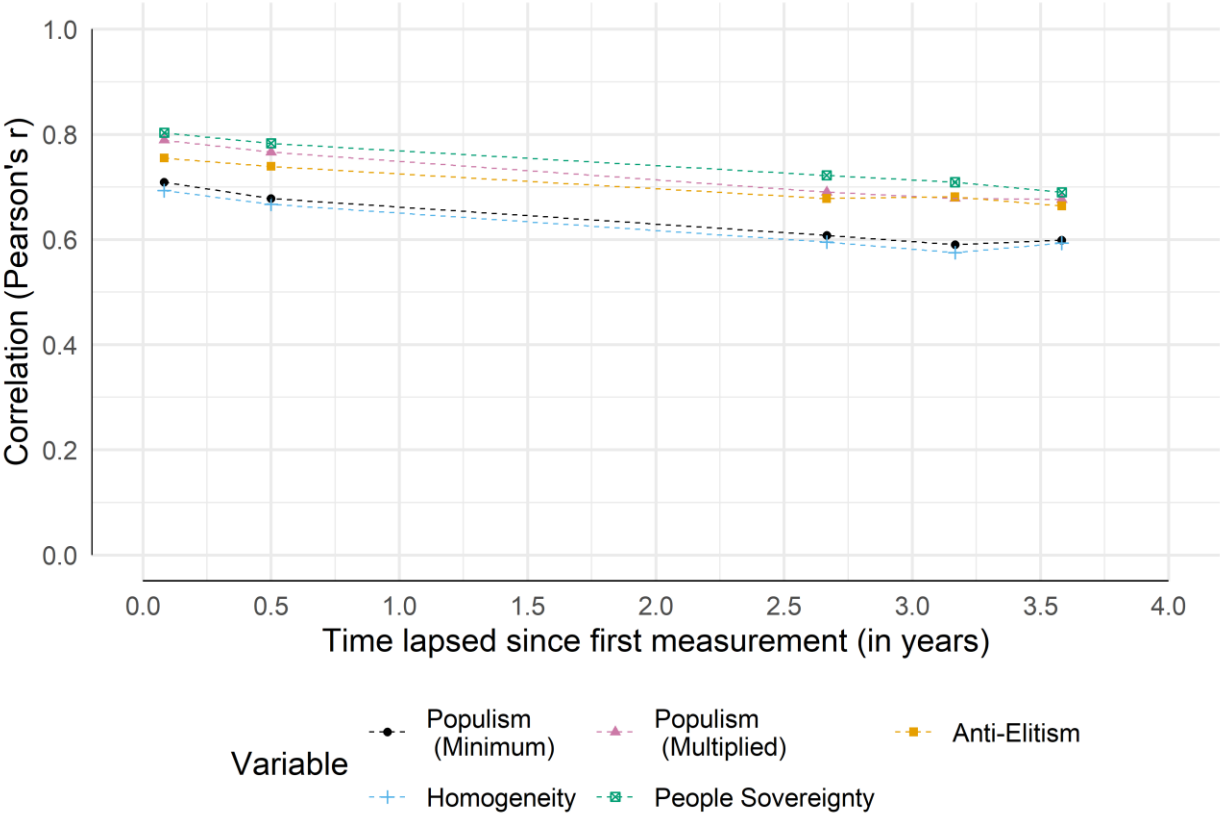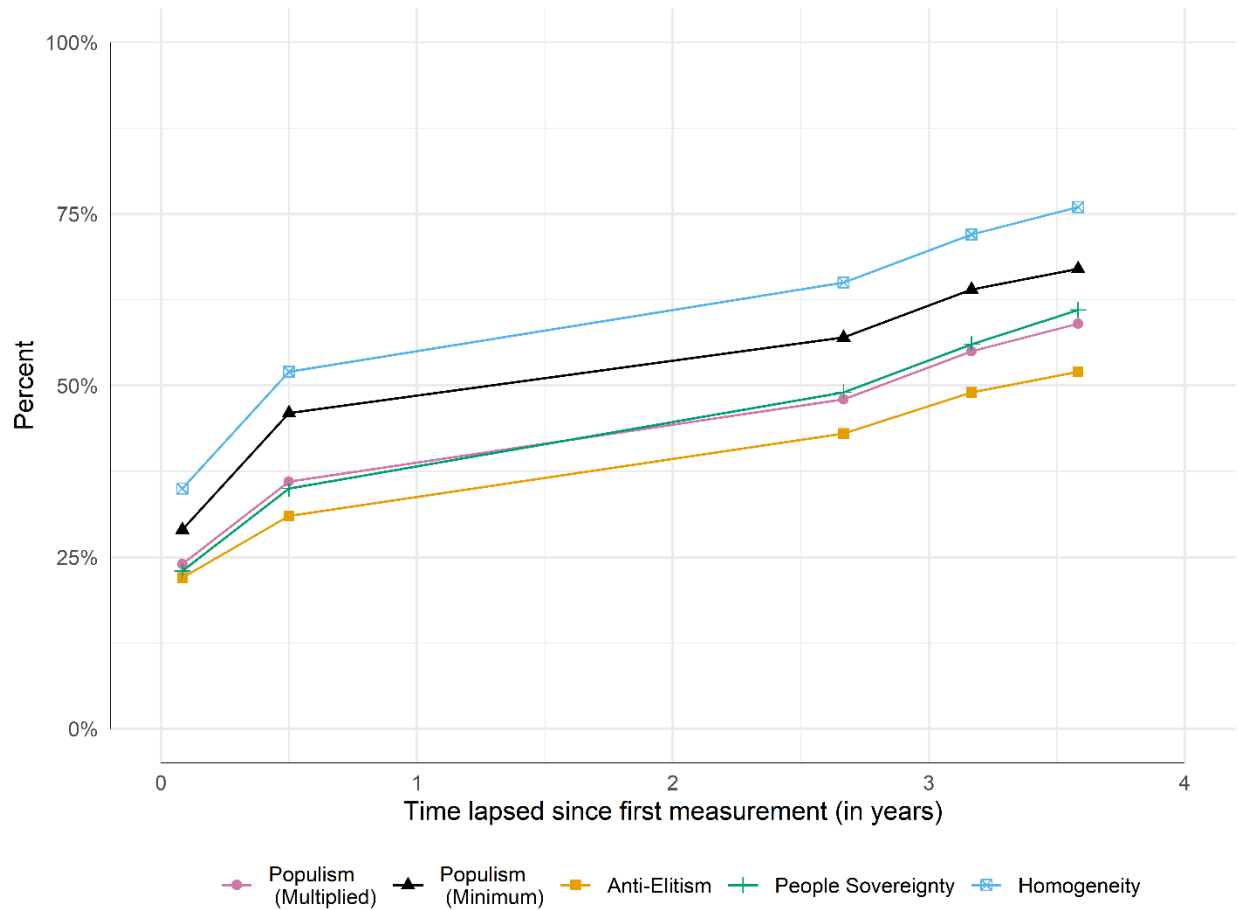
Figure A4.4 Share of respondents with unstable attitudes towards individual dimensions of populism (minimum score and multiplied) and the individual dimensions of populist attitudes over time

Regardless, our main finding remains unchanged. Populist attitudes are neither trait nor widlly fluctuating, however, aggregation rules in multi-dimensional constructs can affect stability when there is heterogeneity in stability at the dimensional level.

Using two-dimensional populism scores

As shown above, the homogeneity dimension in our data is the least stable dimension, both in terms of relative and absolute stability. However, taking out the most volatile dimension and analyzing a two-dimensional populist scale that excludes the homogeneity dimensions does not automatically reduce volatility because the level of stability is dependent in both, the distributions of sentiments and stability across the dimensions. As can be seen in Figure A5.5, the two dimensional construct based on anti-elitism and people sovereignty is more stable relatively

speaking. Yet, Figure A5.6 shows that there is little difference in absolute stability. The reason is that there is considerable within respondent differences in terms of the stability across the two dimensions. For instance, our data shows that of those in the data catgeorized as stable anti-elitists, 48% are unstable people sovereigntists and among respondents coded as stable people sovereigntists, 36% are unstable anti-elitists. Thus, while some of the discrepancies in stability across different aggregated populist scores can be explained by differences between the dimensions, particularly when one dimensions (here homogeneity) presents a strong outlier in terms of sentiment and stability, or aggregation mechanisms (e.g., minimal score versus multiplication), the stability of aggregated scores is also sensitive to underlying distributions and the similarities (or differences) between the dimensions that are included. Note that in the classical approach to scoring where correlations are a key criterium to judge the quality of a scale (Bollen and Lennox 1991), differences are unlikely to be a strong such that they would affect the outcome of this analysis in any particular way. However, since the aggregation of individual dimensions into a single populism score "is not an empirical but and ontological question" (Wuttke et al. 2020, 365), these differences matter.

Figure A5.5 Correlations between measures taken at initial wave and subsequent waves for three-dimensional populist attitude scale (anti-elitism, people's sovereignty, and homogeneity) as shown in main text and two-dimensional populist attitude scale (anti-elitism and people's sovereignty).
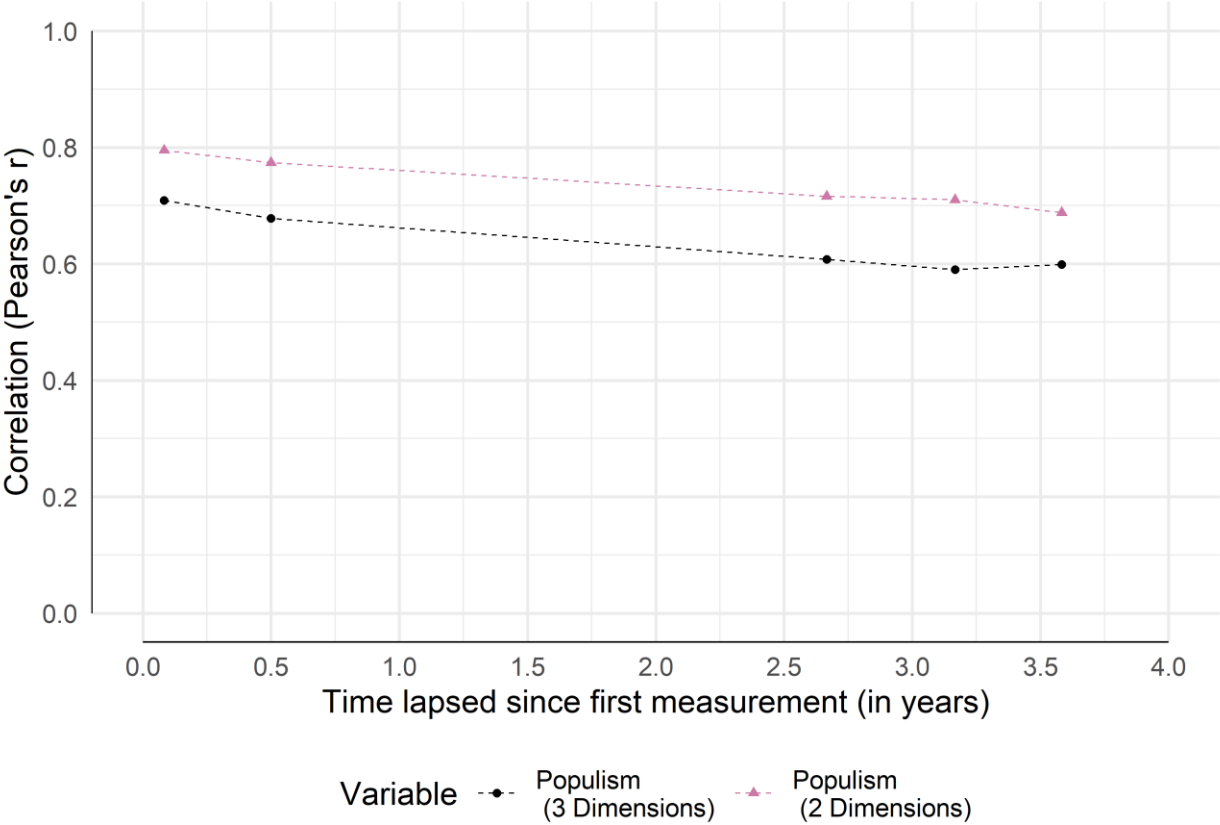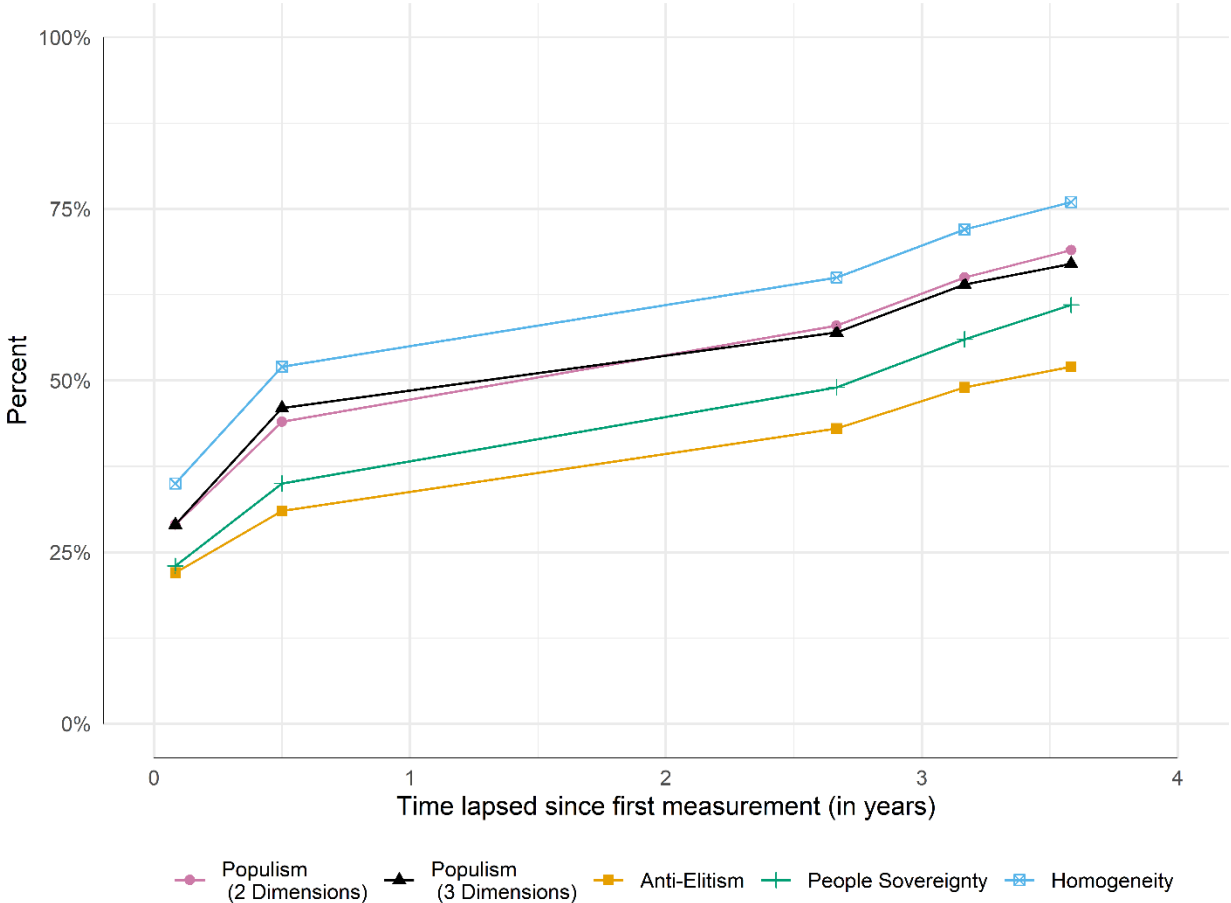
Figure A5.6 Share of respondents with unstable/ambivalent attitudes towards individual dimensions of populism (two and three dimension versions of the scale) and the individual dimensions of populist attitudes over time

Appendix 6 - Additional Correlation Analyses

Correlation with democratic attitudes

In this part of the appendix, we replicate the correlation analyses shown in the main text to highlight the potential substantial impact of setting aside the issue of stable and non stable populist attitude holders. Rather than dichotomizing, as we have done in the main text, Figure A6.1 shows the correlations by the number of times respondents switch opinion boundaries.

Figure A6.1 Correlations between populist attitudes and items on liberal democracy by stability captured as number of times the respondents cross opinion boundaries across the six panel waves
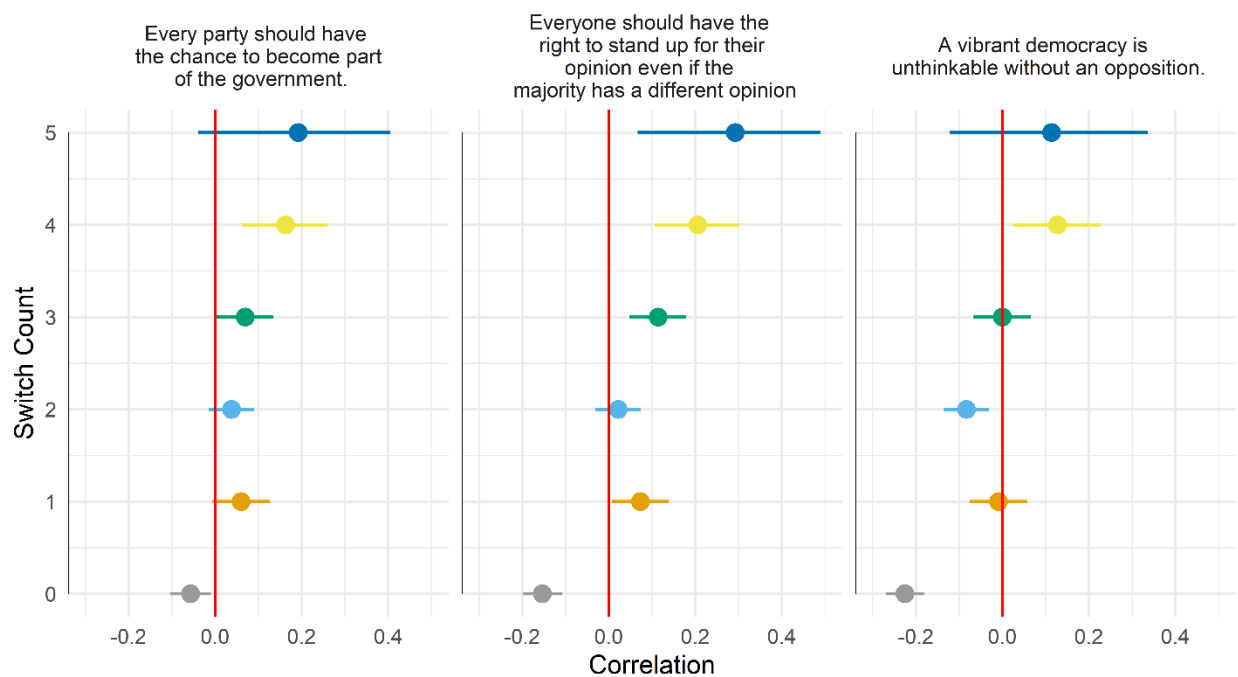
Connection to vote choice

Replicating the analyses using absolute changes

Table A6.1 shows the average absolute changes by populist sentiment and reported vote choice for the AfD. The table shows that, in general, average absolute changes are generally smaller when the self-reported vote choice is in line with populist sentiments. For example, among those who strongly agree with populist attitudes and who voted for the AfD, the average mean change in populist attitudes over time is 0.446. Among those who strongly agree but did not vote for the AfD, the average absolute change is 0.588. Ergo, smaller changes in populist attitudes over time go hand in hand with greater consistency between populist attitude position and political

behaviour. The category "Agree" here is the exception but for both "Disagree" and "Strongly disagree," the pattern holds.

Table A6.1 Reported vote for AfD by support for populist ideas and absolute change in populist attitudes over time

Support for populist ideas	AfD vote?	Mean Absolute Change	SD Absolute Change
Strongly agree	No	0.588	0.599
	Yes	0.446	0.433
Agree	No	0.505	0.521
	Yes	0.565	0.558
Neither agree nor disagree	No	0.375	0.364
	Yes	0.441	0.425
Disagree	No	0.393	0.376
	Yes	0.474	0.449
Strongly disagree	No	0.407	0.406
	Yes	0.526	0.539

Replicating the analyses and using switch count instead of dummy

We replicate the analysis show in the main text, in which we extract of respondents who agree (disagree) with populist ideas and who support the AfD. Rather than distinguishing between stable (zero switches) and unstable (at least one switch), we the count of switches here to group respondents (Table A6.2). The results mirror those in the main text as stable respondents, that is, those who did not switch opinions on populist ideas across the six panel waves, are more likely to behave in accordance with how we would expect populist ideas to affect the support of the AfD, at least in case of respondents who agree with populist ideas. Note that in some cases, the number of respondents for any given category can be small given the detailed raster applied here.

Table A6.2 Reported vote for AfD by agreement with populist ideas and populist intra-individual stability

Support for Populist Ideas	Group/Switch Count	% of Respondents who voted for AfD
Agree	Combined	34%
	0 switches	54%
	1 switch	25%
	2 switches	32%
	3 switches	32%
	4 switches	38%
	5 switches	29%
Disagree	Combined	10%
	0 switches	7%
	1 switch	19%
	2 switches	10%
	3 switches	19%
	4 switches	16%
	5 switches	15%

Replicating the analyses and separating out Die Linke

Table A6.3 Reported vote for AfD and Die Linke by agreement with populist ideas and populist intra-individual stability

Support for Populist Ideas	Group	% of Respondents who voted for AfD	% of Respondents who voted for Die Linke
Agree	Combined	34%	16%
	Unstable	31%	16%
	Stable	54%	15%
Disagree	Combined	10%	17%
	Unstable	14%	17%
	Stable	7%	17%

Appendix 7 - Explorative analysis into determinants of populist attitude stability

In this section, we explore potential factors that can explain populist attitude stability. Specifically, we include the following variables (sources in parentheses):

- Attitude extremity (Howe and Krosnick 2017): Dummy to measure as whether respondents hold extreme populist attitudes (positive or negative) in panel Wave 1 (values smaller than two for respondents who strongly oppose populist ideas and values equal five for respondents who strongly support populist ideas); 0=not extreme, 1=extreme.

- Political interest as proxy for political involvement (Zaller 1992): Level of political interest as measured in panel Wave 1 (see Appendix 1)

- Partisanship as factor that contributes to stability because of motivated reasoning (Lodge and Taber 2000): Dummy to capture whether respondents identify with a party or not, measured in GLES Panel Wave 4/1A; 0=non-partisan, 1=partisan.

- Turnout in 2017 German Federal Election as proxy for political involvement (Zaller 1992): Dummy measured in GLES Panel Wave 8 to capture reported turnout in the election; 0=did not vote, 1=did vote.

- Age (Tyler and Schuller 1991): Continuous variable

- University education as education may lead to more or less stability (c.f. Xu et al. 2020, 1516): Dummy capturing whether respondents hold university degree or whether they do not; 0=no university degree, 1=university degree

- Respondent originates from East Germany to measure different institutional contexts (Svallfors 2010) and populist parties may have had a greater effect on populist attitude levels in the East than the West (e.g. Weisskircher 2020): Dummy capturing whether respondents come former Western German states or not; 0=Respondent comes from former East Germany, 1=Respondent comes from former West Germany.

Regardless of whether we use the grouping of respondents into stable and unstable populist attitude holders (M1 in Table A7.1) or whether we use the number of times respondents cross opinion boundaries (M2 in Table A7.1), the results suggest that attitude extremity and university education might be good starting points for future research to explore the origin of populist attitude stability:

attitude extremity increases likelihood of being in the stable group and reduces number of boundary crossings and university education contributes to stability in populist attitudes. Age and partisanship also come through when considering the number of opinon boundary crossings.

Table 7.1 Results from logistic (M1) and negative binomial (M2) regression models, analyzing potential factors that can explain the stability of populist attitudes

	Dependent variable:	
	Respondent belongs to stable populist attitude group (M1)	Number of opinion boundary crossings (M2)
Political Interest	0.06	-0.02
	(0.04)	(0.02)
Populist Attitude Extremity	1.90***	-1.02***
	(0.12)	(0.07)
Respondent is partisan	-0.18	-0.08*
	(0.09)	(0.04)
Voted in 2017 Federal Election	0.20	-0.11
	(0.16)	(0.06)
Age	-0.01	0.003*
	(0.004)	(0.001)
University Degree	0.47***	-0.15***
	(0.08)	(0.03)
West Germany	0.04	-0.02
	(0.08)	(0.03)
Constant	-1.04***	0.50***
	(0.27)	(0.10)
Observations	3,848	3,848
Log Likelihood	-2,251.02	-6,123.84
AIC.	4,518.05	12,263.68

Note: *p**p***p<0.001; Entries from logistic regression model (M1) and negative binomial model (M2); Standard errors in parentheses.

References

Bollen K and Lennox R (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin* **110** (2), 305–314.

Chen FF (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal* **14** (3), 464–504.

Cheung GW and Rensvold RB (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal* **9** (2), 233–255.

Freeder S, Lenz GS and Turney S (2019). The Importance of Knowing "What Goes with What": Reinterpreting the Evidence on Policy Attitude Stability. *The Journal of Politics* **81** (1), 274–290.

GLES. 2021. GLES Panel 2016-2021, Wellen 1-15.

Hill JL and Kriesi H (2001a). An Extension and Test of Converse's "Black-and-White" Model of Response Stability. *American Journal of Political Science* **95** (5), 397–413.

Hill JL and Kriesi H (2001b). Classification by Opinion-Changing Behavior: A Mixture Model Approach. *Political Analysis* **9** (4), 301–324.

Howe LC and Krosnick JA (2017). Attitude Strength. *Annual Review of Psychology* **68**, 327–351.

Hu L and Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* **6** (1), 1–55.

Kiley K (2021). *Ambivalence is Everywhere: Quantifying Attitude Stability Across Topic Domains*.

Kline RB (2005). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Liu Y *et al.* (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods* **22** (3), 486–506.

Lodge M and Taber C (2000). Three Steps toward a Theory of Motivated Political Reasoning. In Lupia A, McCubbins MD and Popkin SL (eds), *Elements of Reason:* Cognition, Choice, and the Bounds of Rationality. Cambridge University Press, pp. 183–213.

Mackinnon SP and Curtis R (2020). *A Tutorial in Longitudinal Measurement Invariance and Cross-lagged Panel Models Using Lavaan*.

Svallfors S (2010). Policy feedback, generational replacement, and attitudes to state intervention: Eastern and Western Germany, 1990–2006. *European Political Science Review* **2** (01), 119.

Tyler TR and Schuller RA (1991). Aging and attitude change. *Journal of Personality and Social Psychology* **61** (5), 689–697.

Weisskircher M (2020). The Strength of Far-Right AfD in Eastern Germany: The East-West Divide and the Multiple Causes behind 'Populism'. *The Political Quarterly* **91** (3), 614–622.

Wuttke A, Schimpf C and Schoen H (2020). When the Whole Is Greater than the Sum of Its Parts: On the Conceptualization and Measurement of Populist Attitudes and Other Multidimensional Constructs. *American Political Science Review* **114** (2), 356–374.

Xu M *et al.* (2020). Individual Differences in Attitude Consistency Over Time: The Personal Attitude Stability Scale. *Personality and Social Psychology Bulletin* **46** (10), 1507–1519.

Zaller J (1992). *The nature and origins of mass opinion.* Cambridge [England] and New York and NY and USA: Cambridge University Press.