

**ONLINE APPENDICES:**

**THE RISE OF SELF-CENSORSHIP IN AMERICA**

Version 6

**CITATION:**

Sutherland, Joseph L., and James L. Gibson. 2024. “The Rise of Self-Censorship in America.” In Maranto, Robert Anthony, Catherine Salmon, Lee Jussim, and Sally Satel. Editors. *The Free Inquiry Papers*. Washington, D.C.: American Enterprise Institute. Chapter 8.

## **Online Appendix A: The 2020 AmeriSpeak Freedom and Tolerance Survey (FATS)**

Funded and operated by NORC at the University of Chicago, **AmeriSpeak®** is a probability-based panel designed to be representative of the U.S. household population. Randomly selected U.S. households are sampled using area probability and address-based sampling, with a known, non-zero probability of selection from the NORC National Sample Frame. These sampled households are then contacted by U.S. mail, telephone, and field interviewers (face-to-face). The panel provides sample coverage of approximately 97 percent of the U.S. household population. Those excluded from the sample include people with P.O. Box only addresses, some addresses not listed in the U.S.P.S. Delivery Sequence File, and some newly constructed dwellings. While most AmeriSpeak households participate in surveys by web, non-internet households can participate in AmeriSpeak surveys by telephone. Households without conventional internet access but having web access via smartphones are allowed to participate in AmeriSpeak surveys by web. AmeriSpeak panelists participate in NORC studies or studies conducted by NORC on behalf of governmental agencies, academic researchers, and media and commercial organizations.

A general population sample of U.S. adults aged 18 and older was selected from NORC's AmeriSpeak Panel for this study. This survey was offered only in English and was administered on the web and over the phone. Invitations to participate in the survey were initiated on July 1, 2020, and the last interviews were completed on July 24, 2020. In total, NORC collected 1,006 interviews, 950 by web mode and 56 by phone mode.

To encourage study cooperation, NORC sent five email reminders to sampled web-mode respondents. Panelists were offered the cash equivalent of \$5 for completing the study. Interviewed respondents took 29 minutes (median) to complete the survey. The interview was divided into two modules, with the FATS questions asked first. The 29 minutes median is the total length of interview. NORC applied standard cleaning rules to the survey data for quality control by removing responses in the main study interview to questions from non-eligible respondents. These respondents provided responses indicative of speeding through the survey and skipping survey questions. These respondents were not included in the final dataset.

The data are weighted, with various factors going into the construction of the final study weight. These include panel base sampling weights, final panel weights, study-specific base sampling weights, and nonresponse adjusted survey weights.

The weighted AAPOR Response Rate #3 recruitment rate was 23.6 percent, with a weighted household retention rate of 84.8 percent and a survey completion rate of 28.4 percent. A weighted AAPOR Response Rate #3 cumulative response rate of 5.7 percent was achieved. The survey has a margin of error of 4.17 percent, and a design effect of 1.82.

For additional technical information about the sample or the study, email [AmeriSpeak-BD@norc.org](mailto:AmeriSpeak-BD@norc.org) or visit [www.AmeriSpeak.norc.org](http://www.AmeriSpeak.norc.org).

This research was approved by the Washington University in St. Louis Institutional Review Board (IRB). That IRB judged this project to be in the “exempt” category owing to the

fact that participation in the survey was voluntary, no harm was afflicted on the respondents, and no identifiers were connected to the database generated, among other factors.

## Online Appendix B: The Surveys Used in the Time Series

The data included in the time series reported in Figure 1 are derived from: Stouffer's 1954 nationally representative, face-to-face survey<sup>1</sup>; the Nunn, Crockett, and Williams 1973 replication of the Stouffer survey (1978)<sup>2</sup>; the 1987 General Social Survey, as supplemented by the Gibson re-interview<sup>3</sup>; the 2005—2011 Freedom and Tolerance Surveys (FATS), with the 2005 survey being the U.S. part of the 2005 European Social Survey<sup>4</sup>; the 2013—2015 American Panel Surveys (TAPS: <https://wc.wustl.edu/american-panel-survey/>); the 2019 and 2020 AmeriSpeak Freedom and Tolerance Surveys (documented in Online Appendix A, above), and; the 2023 Verasight nationally representative omnibus survey (<https://www.verasight.io/>).

As with any compilation of independent studies, these data are not, strictly speaking, directly comparable owing to different modes of survey administration (and other factors). (For a similar approach to assembling various surveys within a single database, see Enns.<sup>5</sup>) The Stouffer survey, the 1987 GSS/Gibson survey, and the 2005 ESS/US survey were conducted face-to-face. The FATS surveys from 2007 through 2009 were based on RDD samples. The FATS surveys in 2010 and 2011 used RDD samples supplemented with cell phone samples. The TAPS surveys were internet surveys. The 2019 and 2020 surveys were mixed mode surveys (telephone and internet) with internet administration of the instrument very strongly dominant. Care must therefore be taken not to make strict and precise comparisons across the surveys.

The 2023 survey item, gathered through the Verasight survey, introduces “Don’t Know” as a response option (69.9%), in addition to “No, feel less free” (38.1%) and “Yes, do feel as free” (52%). “Don’t Know” responses can indicate ignorance, indecision, or uncertainty about the meaning of the question asked, but it is not clear which of these the respondent intends to communicate without further probing; this probing in itself can induce bias.<sup>6</sup> Survey respondents may also lean in a particular direction but experience cognitive dissonance, or a need to satisfice, in expressing a directional response, making “Don’t Know” a valued “middle response” rather than a “not applicable” non-response.<sup>7</sup> Yes is an easy answer, readily accessible. But no requires

---

<sup>1</sup> Samuel C. Stouffer, *Communism, Conformity and Civil Liberties* (New York: Doubleday, 1955).

<sup>2</sup> Clyde Z. Nunn, Harry J. Crockett, and J. Allen Williams, *Tolerance for Non-Conformity* (San Francisco: Jossey-Bass, 1978).

<sup>3</sup> James L. Gibson, “The Structure of Attitudinal Tolerance in the United States,” *British Journal of Political Science* 19 (1989): 562-570.

<sup>4</sup> James L. Gibson, “Measuring Political Tolerance and General Support for Pro-Civil Liberties Policies: Notes, Evidence, and Cautions,” *Public Opinion Quarterly* 77 (2013): 45-68.

<sup>5</sup> Peter K. Enns, *Incarceration Nation: How the United States Became the Most Punitive Democracy in the World* (New York, NY: Cambridge University Press, 2016).

<sup>6</sup> Sanchez, Maria Elena, and Giovanna Morchio. 1992. “Probing “don’t know” Answers: Effects on Survey Estimates and Variable Relationships.” *Public Opinion Quarterly* 56 (4): 454-474.

<sup>7</sup> Krosnick, Jon A. 1991. “Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys.” *Applied Cognitive Psychology* 5 (3): 213-236.

more thought, and a survey of various contexts. Therefore, a process might include: “I reject yes, but I don’t want to spend the time to see if no is the right answer, so I will just say don’t know.” Therefore, we collapse “Don’t Know” into “No, feel less free” because if respondents don’t know if they feel as free to speak their mind as they used to, then the directionality of the effect is effectively negative.

Nevertheless, the trend in these data is completely obvious and is strong enough to defeat any explanation of the trend in the data based on survey mode.

## Appendix C: The Measurement of Self-Censorship

Several potential threats to the validity of the Stouffer measure of self-censorship can be imagined. Here we address several of those.

### **The Use of a Single-Item Indicator: Why the Stouffer Measure?**

Single-item indicators are, of course, used in a wide variety of research, especially when policy preferences or institutional attitudes are of interest (for example, death penalty, confidence in institutions such as the U.S. Supreme Court, presidential popularity, satisfaction with democracy). When it comes to measures of self-censorship, single-item indicators are much more common than multi-item indicators.

Still, we readily acknowledge that using a single-item indicator to measure self-censorship introduces some amount of measurement error (for a comprehensive set of measures of perceived political freedom see Gibson 1992).<sup>8</sup> Since a central purpose of this research is to measure change in self-censorship—and since the Stouffer survey, conducted during the McCarthy Red Scare, is both one of the oldest and most highly regarded analyses of political intolerance in the United States—we accept the limitations of the measure in order to analyze how responses to the item have evolved over time. This is a tradeoff, to be sure, but it is the only feasible way of getting purchase on the central question of change in public opinion.

### **Defects in the Question Wording?**

Three complaints might be imagined: the question does not explicitly refer to “political speech,” the comparison to “as you used to” is vague, and no response options for “more free” was offered by the question.

#### *Political Speech*

The first issue we address is whether the Stouffer item measures perceived freedom to speak *about politics*. Obviously, the question itself does not mention politics so some might consider that the responses are overly broad in the sense that they conflate and combine political and non-political speech.

We investigated that possibility with a question-wording experiment fielded in a nationally representative survey in September 2022. The survey utilized NORC’s AmeriSpeak Panel, one of the highest quality probability-based surveys available. (The 2019 and 2020 FATS surveys were also conducted via NORC’s AmeriSpeak panel.) Respondents were randomly assigned to hear one of three versions of the “freedom to speak” question. The first version is the original item asked by Stouffer in 1954 (the item reported in the time series in this paper). The second version, simply adds to the Stouffer question the preface “When it comes to politics . . .” The third version was constructed by the Cato Institute and was used in widely noticed 2017/2020 national surveys. The CATO item has traditionally revealed more self-censorship

---

<sup>8</sup> James L. Gibson, “The Political Consequences of Intolerance: Cultural Conformity and Political Freedom,” *American Political Science Review* 86 (1992): 338-356.

than the Stouffer item. We begin by considering the null hypothesis that the version without a reference to politics (“All Speech”) produces the same results as the version with an explicit reference to politics (“Political Speech”). The results are shown in Table C.1.

Clearly, the data indicate that the Stouffer version is understood as asking, *de facto*, about political speech: the differences in the responses between the question delimited to politics and the question silent on what sort of speech is being asked about are absolutely trivial (see the first two columns in the table). We conclude therefore that the Stouffer item is not over-inclusive and is in fact measuring speech relevant to politics.

#### *“as you used to” Comparison*

As for the “as you used to” comparison, we first note that a recent survey in Germany also found it useful to use the past as a referent for assessing contemporary freedom. Menzner and Traunmüller measured self-censorship with the following item: “People like me are no longer allowed to express their opinions freely in public.”<sup>9</sup>

It is also commonplace for survey researchers to ask “as you used to” questions. For example, Pew has asked:

Since Donald Trump's (2016) election, would you say you are paying more, less, or about the same amount of attention to politics as you used to?

And:

Since you started reading newspapers online, are you reading the paper version of the newspaper more often, less often, or about as much as you used to?

CBS News has asked:

In general, because of the economic recession, when you go shopping, are you spending more money on things other than the basic necessities--I don't mean essential food and clothes, but the extras--or less money on these things, or about the same amount of money as you used to?

Fox News has asked:

And what about you--are you personally more grateful for what you have these days or angry you don't have as much as you used to have?

Despite imprecision in knowing exactly what period the respondents are using as the baseline, as a measure of contemporary trajectory, the question should be understood as asking whether “things are getting better or getting worse.” That seems to us to be a quite valid approach.

---

<sup>9</sup> Jan Menzner and Richard Traunmüller, “Subjective Freedom of Speech: Why Do Citizens Think They Cannot Speak Freely?” *Politische Vierteljahresschrift* (2022): 1-27.

TABLE C.1.

*Alternative Measures of Perceived Freedom to Speak, 2022*

Response	<i>All Speech</i>	<i>Political Speech</i>	<i>Prevents Me from Saying</i>
Free to Speak	42.9	41.1	32.5
Not Free to Speak	57.1	58.9	67.5
Total	100.0	100.0	100.0
N	630	345	346

Source: NORC AmeriSpeak 2022

Notes:

Table entries are percentages (except for N).

*All Speech*: What about you personally? Do you or don't you feel as free to speak your mind as you used to?

Yes, do feel as free

No, feel less free

*Political Speech*: What about you personally? When it comes to politics, do you or don't you feel as free to speak your mind as you used to?

Yes, do feel as free

No, feel less free

*Prevents Me from Saying*: The political climate these days prevents me from saying things I believe because others might find them offensive.

Strongly agree

Somewhat agree

Somewhat disagree

Strongly disagree

Source: NORC AmeriSpeak 2022



### *No “more free” Option Offered*

We are not much concerned about this issue. First, we expect that not many would say that they are “more free” (especially in light of the Cato findings that we cite). Second, the question in essence asks, “are you less free, yes or no?” We readily concede that those who say they are not less free include two types of people: those for whom there has been no change in their perceived level of freedom and those who see themselves as having more freedom. Our analysis focuses on describing and understanding how the replies of “less free,” which we contend are *not* ambivalent, have changed over time and how they vary across individuals and communities. So, the heterogeneity in the “not less free” category is not of much concern to us.

### **Convergent Validity**

How do the responses to the Stouffer item stack up against other measures of self-censorship?

#### *The Cato Measure*

The Cato measure of self-censorship—“The political climate these days prevents me from saying things I believe because others might find them offensive”—seems not to suffer from any obvious question-wording infirmities. It refers to speech in a political context, identifies the source of constraints on speech as being external to the respondent, and is a contemporaneous measure (not requiring any comparison to the past). As noted, a portion of the AmeriSpeak sample was assigned (randomly) to respond to this item.

The data in Table C.1 (above) reveal that in 2022 the Stouffer item provides more conservative estimates of self-censorship than the Cato measure (which was also true in the 2020 survey), with the Stouffer measure registering about 10 percentage points less self-censorship than the Cato measure. In its 2017 survey, Cato found that 58 percent of Americans engaged in self-censorship; in its 2020 survey, the number had risen to 62 percent. Still, the alternative measures of self-censorship all support that conclusion that a large proportion of the American people do not believe they enjoy unrestricted freedom of speech.

Finally, it should be noted that the 2022 data indicate for both the Stouffer version and the Cato question considerably more self-censorship in 2022 than in 2020.

#### *The 1987 General Social Survey*

We also endeavored to validate the Stouffer measure with some items that were first used in a 1987 GSS reinterview survey (reported by Gibson 1993) pertaining to reluctance to speak one’s mind. Figure C.1 reports the relationship between the responses in the 2022 survey to the Stouffer item and these measures of hesitancy about speaking. The statistic reported in the graph is the difference in the percentage of those who self-censor according to the Stouffer measure who express reluctance to speak out for various reasons minus the percentage for those who do not self-censor. Note that our expectations are that the Stouffer responses will be related to perceived external constraints on freedom speech but not to perceived internal constraints.

The figure reports interesting and expected relationships and non-relationships. Regarding the former, those who currently feel less free to speak their minds are also much more likely than those who do feel free to express their views to assert that talking about politics

creates enemies, that they worry about what people might think of them, that they do not like arguments, and that they expect that others would judge their views to be strange. All of these seems to be beliefs that reinforce or explain their perceptions of a lack of freedom by references to external constraints.

At the time, however, there is little or no difference between those feeling unfree and those not feeling unfree on beliefs that they are insufficiently knowledgeable to speak, that their associates do not care about politics, or that they hold similar views to their associates, and for other reasons as well. In general, those feeling less free identify the expected external constraints on their freedom to speak and fail to identify the expected internal constraints on their freedom, providing at least some confidence in the validity of the Stouffer question.

Our 2022 validation survey also repeated a question from the 1987 GSS on the degree to which people are worried about expressing their political views to others. The question asked about being worried to express their views in a variety of contexts, ranging from publicly in their community to their immediate family. The 1987/2022 question read:

How worried are you about expressing your political views to . . . ?

- Members of your immediate family
- Your close friends
- Your co-workers
- Members of organizations to which you belong
- Publicly, in your community
- Your representatives in the government

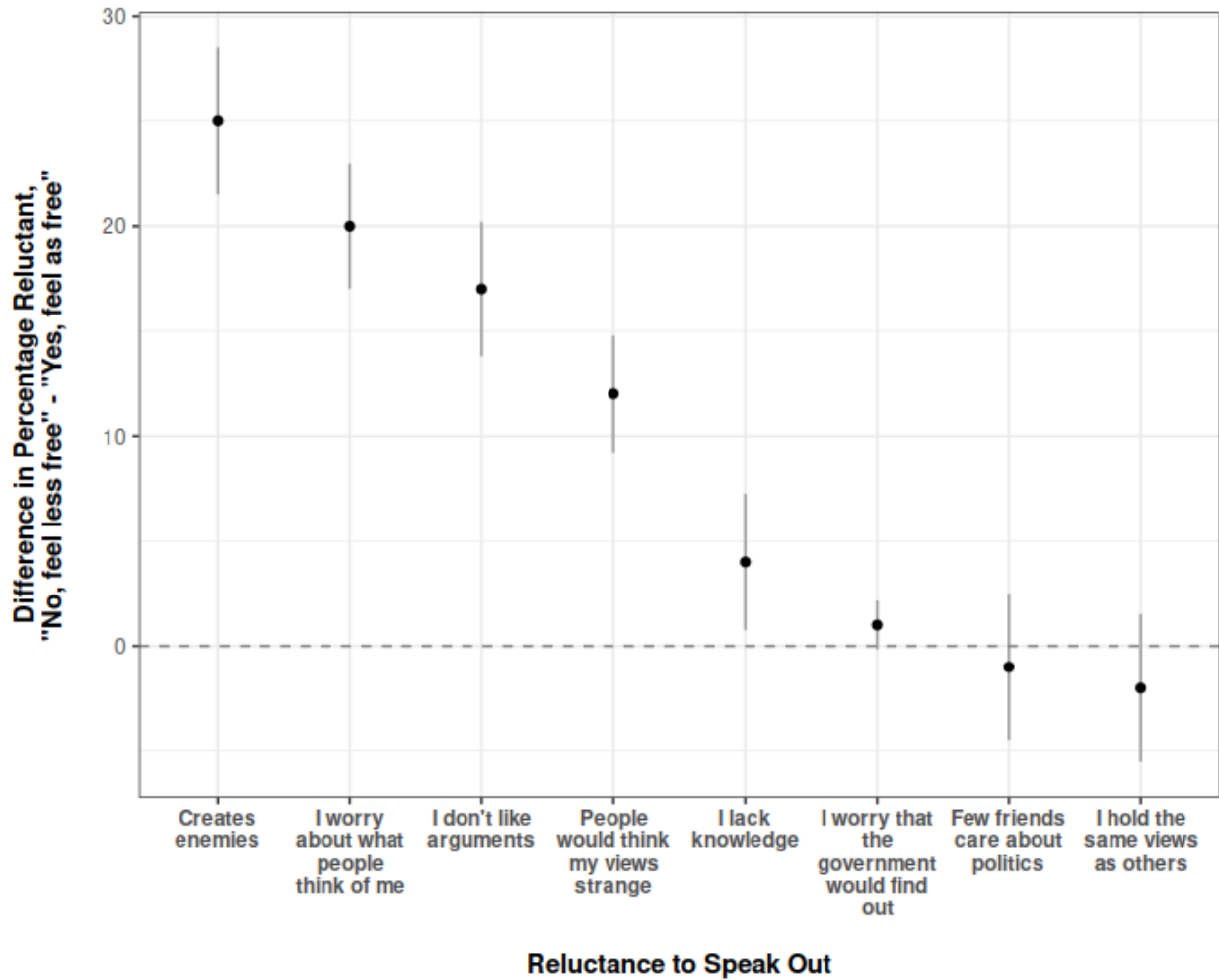
The response set was:

- You worry quite a bit about what they will think of your views
- You worry only some about what they will think of your views
- You don't worry very much about what they will think of your views
- You don't worry at all about what they will think of your views

We created an index of the extent of being worried across the six contexts by scoring those worried "quite a bit" and "only some" as worried and those who "don't worry very much" and "don't worry at all" as not worried. Regarding the Stouffer item, those feeling free to express themselves had an average of 1.1 contexts in which they were worried, whereas those feeling less free worried about expressing themselves, on average, in 1.9 contexts. The difference between the two Stouffer groups is statistically significant at  $p < .001$  (and the same results are produced using different summary measures of worries).

FIGURE C.1.

*Correlates of Self-Censorship, 2022*



Notes:

The table entries are the differences in the percentages of people who self-censor who regard the statement about reluctance to speak as true as pertains to them minus the percentage of those who do not self-censor.

The question read:

Some people have told us that they are occasionally reluctant to talk about politics with their families and friends. I would like to read you several statements and ask if they are true or false as they apply to you.

*Creates enemies*<sup>\*\*\*</sup>: I am sometimes reluctant to talk about politics because it creates enemies.  
*I lack knowledge*: I am sometimes reluctant to talk about politics because I lack the information and knowledge to do so.

*People would think my views strange*<sup>\*\*\*</sup>: because people would think my political views were strange.

*Few friends care about politics*: because few of my friends and family care about politics.

*I hold the same views as others*: because I usually hold the same political views as those I am around.

*I don't like arguments*<sup>\*\*\*</sup>: I am sometimes reluctant to talk about politics because I don't like arguments.

*I worry about what people think of me*<sup>\*\*\*</sup>: because I worry about what people would think of me.

*I worry that the government would find out*: because I worry that the government might find out about me.

<sup>\*\*\*</sup> The difference in proportions between those feeling free and those not feeling free is statistically significant at  $p \leq .001$

95 percent confidence intervals are shown for each difference of percentages.

Source: NORC AmeriSpeak 2022

---

### **Reliability**

Finally, the reliability of the responses to single-item indicators is more difficult to assess. In such circumstances, test/re-test reliability is the most useful approach. In the 2005 FATS survey, a subsample of respondents was reinterviewed twice. The original interview was face-to-face; the follow-up interviews were conducted by telephone. Of the 1,001 respondents in the  $t_1$  survey, 257 were reinterviewed at  $t_3$ . The  $t_3$  interview took place approximately one year after the  $t_1$  interview. The Stouffer question was asked at  $t_1$  and  $t_3$  (but not at  $t_2$ ). For 64 percent of the respondents, the same answer was given at  $t_1$  and  $t_3$ . A large majority of those saying they felt free at  $t_1$  also said they felt free at  $t_3$ . For those reporting not feeling free at  $t_1$ , about one-half said they did not feel free at  $t_3$ . These results are somewhat challenging to assess because such a small percentage of the  $t_1$  respondents completed the  $t_3$  interview (in part because completing the  $t_2$  interview was a requirement for being reinterviewed at  $t_3$ ), because of the shift in interview mode, and because a year elapsed between the  $t_1$  and  $t_3$  interviews. Still, that approaching two-thirds of the respondents gave the same response at both interviews suggests a reasonable degree of reliability in the measure.

Returning to the time-series data, we also observe that the pattern of change over time in levels of self-censorship seems relatively smooth, with considerable similarity in the results from surveys conducted in close temporal proximity to each other. Such a pattern does not seem

compatible with the hypothesis that the Stouffer measure is contaminated with a great deal of random error (a reliability matter).

### **Are Responses Driven by Contexts?**

Every analyst much face the possibility that the context of questions and answers changes over time (even over fairly short periods of time). Most long-term repeated cross-sectional surveys (for example, the GSS, ANES) are biased in favor of not updating question wording even when the wording seems to have become less apposite to current conditions (for example, GSS's fixed-group tolerance questions). For instance, the question "should a communist be allowed to give a speech" no doubt is embedded in different contexts when asked in 1954 and when asked in 2022 (as is true of all the groups about which the GSS asks its tolerance questions.<sup>10</sup> The same could be said about longitudinal analyses of racial attitudes, or, for that matter, virtually any issue (for example, health care, abortion, affirmative action, the death penalty). Even the composition of "liberals" and "conservatives" has undoubtedly changed over time.

Our view is that the description of public opinion and the analysis of the etiology of public opinion must be treated as different questions. We have no doubt that one possible explanation of the rise of self-censorship in the United States revolves around the growth of social media usage (although one should be careful not to over-estimate the importance of social media, especially for middle-aged and older Americans). Widely used available time-series measures—such as Stimson's measures of "mood" or Enns' measures of punitiveness—face the same issues.<sup>11</sup> Indeed, the shifting definitions of concepts are one reason why it is important to use the same question wording over time.

### **Conclusions**

Generally, these various empirical findings from the 2022 survey tend to support the view that self-censorship is being validly and reliably measured by the Stouffer item. We certainly do not maintain that the Stouffer item is a perfect summary measure of self-censorship, but we do contend that there is sufficient evidence of the psychometric properties of the item to warrant using the measure in this paper and to support the substantive conclusions we draw.

Our purposes in this paper are to describe change in levels of self-censorship in the United States, to provide some highly exploratory correlates (and non-correlates) of that change, to point to unanswered questions requiring further research, and to suggest a theoretical framework for such additional investigations. Focusing on the Stouffer item therefore seems to be quite reasonable. In the final analysis, we contend that our evidence of change over time is sufficiently dramatic that it most likely cannot be attributed to temporal frailties in the survey question.

---

<sup>10</sup> Dennis Chong and Morris Levy, "Competing Norms of Free Expression and Political Tolerance," *Social Research* 85, no. 1 (2018): 197-227.

<sup>11</sup> James A. Stimson, *Public Opinion in America: Moods, Cycles, and Swings*. 2nd Edition (New York: Routledge, 2019); Enns, *Incarceration Nation*.

## Online Appendix D: The Time-Series Database

As we have noted in the text, our collection of various yearly time series does not overlap very well when we match the data by the exact year. For example, ideological affective polarization and partisan affective polarization are simultaneously available in only 15 years (for which the correlation is .70). Because our objective in this analysis is nothing more than to offer a rough comparison of trends in different time series, we have created a full data series in which scores between years are linearly interpolated.<sup>12</sup> This database was used to create Figures 2 and 3 in the text. Across the time period from 1954 to 2020, the correlations of the variables shown in the graphs are (number of cases available in parentheses):

Self-Censorship	1.00				
Civil Liberties Support	-.73 (45)	1.00			
Ideological Affective Polarization	.91 (48)	-.65 (48)	1.00		
Partisan Affective Polarization	.86 (43)	-.82 (43)	.77 (43)	1.00	
Year	.85 (48)	-.95 (45)	.74 (48)	.94 (43)	1.00

Some might object to this database because the yearly coverage of surveys in the period prior to 2005 is sparse indeed. We therefore report the correlation matrix for these variables for the period from 2005 to 2020 (N = 16).

Self-Censorship	1.00				
Civil Liberties Support	-.92	1.00			
Ideological Affective Polarization	.89	-.97	1.00		
Partisan Affective Polarization	.70	-.77	.87	1.00	
Year	.87	-.93	.92	.73	1.00

Because the conclusions that we seek to draw from this analysis are minimalist, simply related to trends over time, and because these correlations are so strong and unambiguous and so little affected by the time period to which they are constrained, we have confidence that our figures do not mis-represent our data.

---

<sup>12</sup> Avram Sidi, *Practical Extrapolation Methods: Theory and Applications* (New York: Cambridge University Press, 2003).