The Rise of Affective Polarization: Is It Who We Are, or What We Think?

John Konicki*†

August 17, 2026

Conditionally Accepted at the American Journal of Political Science

Abstract

Scholars express alarm at rising affective polarization. Many blame rising demographic sorting, arguing voters' social characteristics ("who we are") have increasingly aligned with their party. Others point to partisan divides around ideology, policy preferences, and group sentiments ("what we think"). Unfortunately, past research lacks comparable over-time metrics of demographic and attitudinal sorting—and often conflates the two. Using measures that are comparable across time and domains, I find that demographic sorting has not meaningfully increased, nor has its relationship with polarization strengthened. Even with generous causal assumptions, demographic sorting can only explain about 2.2% of the rise in affective polarization. By contrast, attitudinal sorting has increased substantially and can account for a substantial share of polarization's rise, with panel and meta-analyses suggesting that attitudes fuel polarization while polarization, in turn, reshapes attitudes. Thus, the rise of affective polarization is bound up in what we think, not who we are.

Word count: 9,994 words

* PhD Candidate, University of California, Berkeley, john_konicki@berkeley.edu

† I thank Sarah Anzia, Kirk Bansak, David Broockman, Ryan Brutger, Christoffer Dausgaard, Pia Deshpande, Nic Dias, Yph Lelkes, Gabe Lenz, Jeremy Levy, Jack Maedgen, Tyler Reny, Eric Schickler, Sean Westwood, Martha Wilfahrt, Chris Wlezien, attendees at the 2025 Polarization Research Lab's annual meeting, and three anonymous reviewers for their helpful feedback. I also thank Lilliana Mason for providing replication materials.

"Today, the parties are sharply split across racial, religious, geographic, cultural, and psychological lines. There are many, many powerful identities lurking in that list, and they are fusing together, stacking atop one another, so a conflict or threat that activates one activates all."
– Ezra Klein, Why We're Polarized (2020, p. 136)

INTRODUCTION

Scholars have expressed grave concern about the rise in affective polarization in America, leading many non-academic observers to take notice as well (Druckman et al. 2024; Finkel et al. 2020; French 2020; Iyengar et al. 2019; Klein 2020; Sasse 2018). One influential explanation attributes this rise to demographic sorting: the allegedly stronger alignment between party identification and *who we are*—social group memberships such as race, religion, and region. For example, Mason (2018) warns of an "American identity crisis" that "emerges when partisan identities fall into alignment with other social identities, stoking our intolerance of each other to levels that are unsupported by our degrees of political disagreement" (p. 63). By contrast, other research emphasizes those disagreements by linking affective polarization to attitudinal sorting. This work shows how partisans have become more divided by *what we think*—ideology, policy preferences, and sentiments towards social groups (Bougher 2017; Levendusky 2009; Robison and Moskowitz 2019).

No prior work has tested the relative merits of these accounts. Past studies have considered these trends separately or conflated them, and researchers lack comparable over-time measures of demographic and attitudinal sorting. This is unfortunate—we must understand the causes of America's growing political divides if we wish to address them. To this end, I develop methodologically comparable measures of demographic and attitudinal sorting based on random forest prediction models, enabling a direct assessment of how much each form of sorting could have contributed to the rise of affective polarization.

I propose that, for any form of partisan sorting to cause an increase in polarization over time, one of two conditions must be met: Either sorting levels must increase over time (while affecting polarization), or sorting's effect on polarization must strengthen over time. This paper finds, first, that demographic sorting has not significantly increased since 1952—and increased only slightly since 1972. Second, the relationship between demographic sorting and affective polarization, though significant, has not strengthened with time. Thus, neither necessary condition is met: Demographic sorting cannot explain the rise in affective polarization.

Attitudinal sorting, by contrast, may play an important role. I document substantial increases in sorting based on ideology, policy preferences, group sentiments, economic attitudes, cultural attitudes, and racial attitudes. All forms of attitudinal sorting predict affective polarization, and racial attitude sorting has likely become more predictive over time.

Based on these findings—and using the most generous assumptions possible regarding the causal relationship between sorting and affective polarization—I estimate demographic sorting can only directly explain 2.2% of the increase in polarization since 1980 (or 6.1% indirectly), whereas attitudinal sorting can explain more than 70%. Within attitudinal sorting, policy preferences appear to explain more than symbolic ideology or group sentiment sorting. However, cultural and racial attitudes can each explain more than economic attitudes, suggesting group-based attitudes have played a substantial role in rising polarization, particularly in recent years.

Next, I reanalyze prior works which claim demographic sorting causes affective polarization. This research incorporates identity importance alongside demographic group memberships, but it also includes non-demographic measures of ideological and partisan identity.

Reanalyses show these political identities drive most of the apparent relationship between social sorting and polarization, with a substantially smaller role for demographic identity strength.

Finally, I examine the extent to which attitudinal sorting causes affective polarization with panel analyses and a reanalysis of past randomized experiments. Although no single approach is definitive, together these studies suggest that attitudinal sorting and affective polarization reinforce each other.

Ultimately, demographic sorting is a stable predictor of affective polarization, but it cannot explain the steep rise in polarization over time. Contemporary polarization seems far more defined by substantive disagreements than by a decline in cross-cutting social cleavages, although these disagreements increasingly focus on identity-inflected domains. Thus, the roots of rising affective polarization appear to be in "what we think" rather than "who we are." Our divisions appear more substantive than conventional wisdom often suggests.

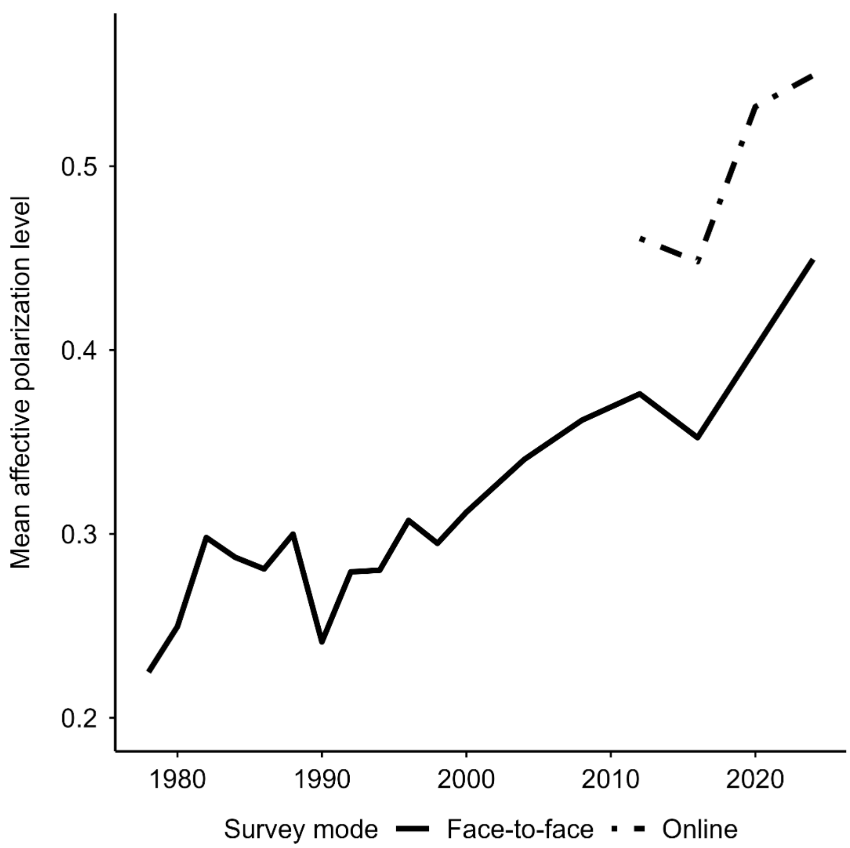
BACKGROUND

Scholars have long recognized that common ground between political coalitions tamps down conflict. In mid-twentieth-century America, as Campbell et al. (1960) observed, there were "large representations of both parties at virtually all social and occupational levels", and majorities of voters did not perceive differences in party positions on policy issues (p. 150). Such overlap and ambiguity were seen as depolarizing forces: Scholars credited the relative stability of US politics to these cross-cutting pressures (Agar 1950; Dahl 1981), and early public opinion research linked them to political ambivalence (Berelson et al. 1954; Campbell et al. 1960; Lazarsfeld et al. 1944; cf. Horan 1971).

Recent scholarship has drawn on these insights to explain rising affective polarization with theories of "partisan sorting." Such sorting, at the individual level, means having traits that match one's co-partisans and distinguish one from out-partisans. Sorting is therefore a combination of individual traits and group distributions of those traits. This means a partisan can become increasingly sorted in three ways:

1. Her traits change in ways that better match her co-partisans
2. She switches parties and has more in common with her new co-partisans
3. She stays the same while her co-partisans increasingly resemble her

Figure 1: Affective polarization levels over time



Feeling thermometer polarization levels in the ANES, over time. Polarization is measured as the difference between in-party and out-party thermometer ratings.

Demographic Sorting and Affective Polarization

According to one prominent theory, the parties have undergone decades of demographic sorting. The parties once had substantial demographic overlap but became increasingly internally homogeneous in response to developments such as the 1964 Civil Rights Act and the rise of the Moral Majority. Thus, as Mason (2018) argues, “A single vote can now indicate a person’s partisan preferences *as well as* his or her religion, race, ethnicity, gender, neighborhood, and favorite grocery store” (p. 14). Because the parties became socially distinct, the typical voter now has fewer demographic connections to the opposite party, leading to a rise in affective polarization. As Table 1 shows, this account has become central to scholars’ understanding of modern American polarization.

The theory is consistent with decades of research showing that social groups shape Americans’ vote choices and partisan identity (Achen and Bartels 2017; Campbell et al. 1960; Green et al. 2002). Furthermore, stark partisan cleavages have emerged around race (Barreto and Segura 2014; White and Laird 2020), religion (Jones 2016; Margolis 2018), urbanicity (Jacobs and Shea 2023; Rodden 2019), and education (Grossmann and Hopkins 2024; Marble 2024; Zingher 2022). Partisans are aware of these divides, readily classifying groups as Republican or Democratic and demographically stereotyping the parties (Ahler and Sood 2018; Kane et al. 2021; Rothschild et al. 2019). These trends suggest a fusion of partisanship and demographics in voters’ minds (Elder and O’Brian 2022; Westwood and Peterson 2022; Zhirkov and Valentino 2022).

However, there are grounds for skepticism about whether demographic sorting can explain an *increase* in affective polarization. For this, it is necessary (though not sufficient) for at least one of two conditions to be true:

1. Demographic sorting must have increased in the electorate (and cause polarization to some degree).
2. Demographic sorting’s impact on polarization must have strengthened over time.

Table 1: Academic perspectives on demographic sorting and polarization

Perspective	Source
"[W]hen a nation changes from one made up of many cross-cutting identities to one built on socially segregated parties, the result is an electorate that is, on average, more angry, excitable, and active than it was before the social shift."	Mason (2018, p. 141)
"As partisan and ideological identities became increasingly aligned, other salient social identities, including race and religion, also converged with partisanship. White evangelicals, for instance, are overwhelmingly Republican today, and African Americans overwhelmingly identify as Democrats. This decline of cross-cutting identities is at the root of affective polarization."	Iyengar et al. (2019, p. 134)
"[A]s America's partisan camps become increasingly homogeneous, cross-cutting identity groups, once seen as a staple of partisan compromise, are on the verge of extinction."	Weber and Klar (2019, p. 218)
"The parties also have sorted along racial, religious, educational, and geographic lines. Although far from absolute, such alignment of ideological identities and demography transforms political orientation into a mega-identity that renders opposing partisans different from, even incomprehensible to, one another."	Finkel et al. (2020, p. 534)
"As an outcome of social sorting, partisans are now more distinct from each other demographically than in the past, with fewer cross-cutting identities to stabilize interparty relations. Growing partisan social dissimilarity has exacerbated an 'us versus them' mentality resulting in greater negative affect toward partisan rivals."	Huddy and Yair (2021, p. 292)
"Republicans and Democrats have grown increasingly socially divided among racial, religious, and other lines in recent decades. These social divisions between the parties have reduced the prevalence of 'cross-cutting' identities that tend to dampen intergroup partisan conflict. It is easier to view opposing partisans with extreme hostility when we can find nothing in common with them than when we can relate."	Kalmoe and Mason (2022, p. 78)
"[T]he number of people with cross-cutting social groups has drastically declined as a result of partisan sorting. … When social groups do not overlap, outgroup members simply cannot understand each other and begin to see each out-group extremely negatively."	Dyck and Pearson-Merkowitz (2023, p. 28)
"[T]he parties have become socially distinctive and homogenous, leading people to harbor more animus toward the other party, which looks less and less like them."	Druckman et al. (2024, p. 11)
"Partisan divisions more frequently fell along the lines of other social boundaries such as race, religion, generation, and place of residence; as fewer citizens held identities that cut across these categories (e.g. a born-again Christian Democrat; a big-city Republican), they became more likely to perceive their own side as representing 'us' and the other party as 'them.'"	Grossmann and Hopkins (2024, p. 7)
"Today, cleavages have 'stacked' in a manner that promotes partisan animosity. Stark divisions among voters—race, ideology, geography, religion, education—align with party. … [T]his alignment—and the associated degradation of crosscutting social ties—has turned partisanship into a kind of 'mega identity.' … When a range of social identities all push in a single direction it becomes much easier to see one's opponents as socially distant and perhaps deserving of hatred."	Pierson and Schickler (2024, p. 17)

Surprisingly, prior research has not established that either condition holds. And there are several reasons to doubt that they do.

First, topline comparisons show that the parties are more demographically similar than stereotypes imply. As of 2024, both parties still have white and non-college-educated majorities, Christian pluralities, and similar median incomes.[1] Second, demographics remain weak predictors of party ID, especially compared to ideology (Abramowitz 2011; Hare and Kutsuris 2023) and affect toward social groups (Kane et al. 2021). Indeed, demographics' predictive power has been stable for party ID and presidential vote since the 1950s (Kim and Zilinsky 2024; see also Calvo et al. 2024). Although binary predictions cannot provide a continuous measure of demographic cross-pressures—and therefore do not distinguish between moderately and extremely sorted voters—these findings nevertheless raise questions about whether demographic sorting has actually increased.

Third, demographics are significantly more stable than attitudes, ensuring one common mechanism of partisan sorting—changing one's traits to match one's party—is relatively unlikely for demographic sorting.[2] Fourth, existing research shows demographic sorting predicts higher affective polarization but does not investigate changes over time—and therefore does not distinguish between demographic sorting as a driver of surging polarization versus a significant but stable contributor (Mason 2018; Harteveld 2021; Guntermann 2022; Jin et al. 2025).

Finally, the most prominent research on demographic sorting uses measures of "sociopartisan" sorting that rely on more than demographic categories (Mason 2018; Mason and

[1]In the 2024 ANES, Republicans (Democrats) are 78% (53%) white, 74% (54%) have no bachelor's degree, 53% (45%) are Christian, and they have a median income of $90k–$100k ($100k–$110k).
[2]However, demographics are far from immutable, and they are sometimes endogenous to party ID (Agadjanian and Lacy 2021; Egan 2020; Margolis 2018).

Wronski 2018). These measures also incorporate demographic identity strength[3] and treat political identities (such as ideology) as additional social groups. As I discuss in Analysis #4, these constructs are combined into a single additive measure, which prevents comparisons between the contributions of demographic categories, demographic identity strength, and non-demographic groups.

Attitudinal Sorting and Affective Polarization

Attitudinal sorting may also increase affective polarization. According to Bougher's (2017) "belief congruence theory" of polarization, voters understand the parties as groups that differ based on beliefs, not just demographics. Thus, if the distributions of attitudes become more homogeneous within each party—and more distinct between them—partisans will encounter fewer cross-pressures, viewing their own party as more consistently aligned with their beliefs and the out-party as more opposed. This should make voters more hostile to their political opponents.

Many attitudinal divides have emerged in recent decades. Most prominently, scholars agree that ideological and policy sorting have increased, even if they disagree on whether attitudes have become more extreme (Abramowitz and Saunders 2008; Baldassarri and Gelman 2008; Campbell 2016; Fiorina et al. 2005; Levendusky 2009). Voters are aware of these partisan differences—often to the point of overstating them (Levendusky and Malhotra 2016b; Druckman et al. 2022; Dias et al. 2025). Furthermore, both ideological and policy sorting are predictive of affective polarization and related constructs (Bougher 2017; Hillygus and Shields 2008; Levendusky 2009; Webster and Abramowitz 2017).

[3]Throughout this paper, "demographic sorting" refers to sorting based on demographic categories, whereas "demographic identity" refers to self-reported importance of group affiliations to respondents' identities (see Brubaker 2002; Huddy 2001; Jardina 2019).

Scholars debate whether policy disagreements reflect sincere differences or just cue partisan team membership, but both perspectives are compatible with the attitudinal sorting hypothesis: Sorting enhances the magnitude of real disagreements *and* turns policy preferences into reliable signals of partisanship (Orr and Huber 2020; Dias and Lelkes 2022; Orr et al. 2023). A similar debate persists over the nature of symbolic ideology: Does it represent genuine policy preferences, a social group reflective of values and lifestyle, or merely the strongest signal of partisanship? (Broockman 2016; Ellis and Stimson 2012; Kollman and Jackson 2026; Simas 2023; Yeung and Quek 2025) No matter which theory is true, sorting along this dimension magnifies partisan differences.

Another form of sorting centers on feelings toward social groups—what Kane et al. (2021) term "group sentiments"—ranging from labor unions to Christians to immigrants.[4] Group sentiments have long been seen as strong forces in American politics because they provide powerful yet accessible heuristics to voters (Campbell et al. 1960; Conover 1988; Sears 1993). These attitudes durably predict partisan affect, and the group-sentiment divide between the parties has widened over time (Kane et al. 2021; Mason et al. 2021; Miller and Wlezien 1993; Robison and Moskowitz 2019).

Finally, other research tracks divides by attitude *domain*, focusing on partisan differences in economic, cultural, and racial views. Earlier work indicated the parties were more divided by economics than culture or race (Abramowitz 1994; Ansolabehere et al. 2006; Bartels 2006), with economic attitudes best predicting affective polarization (Iyengar et al. 2012; Webster and Abramowitz 2017). However, partisan conflict has increasingly realigned along cultural and racial debates (Marble 2024; Sides et al. 2022).

[4]Kane et al. (2021) differentiate between group membership and group sentiments, noting "group membership alone is not a necessary condition for group sentiments to shape partisanship" (pp. 1784–1786).

Attitudinal sorting's ability to drive affective polarization is by no means obvious. Some scholars are skeptical of attitudes' explanatory power beyond symbolic ideology, arguing Americans are polarized primarily by their identities and otherwise "disrespectfully agree" on substantive matters (e.g., Mason 2015). Others doubt that ideology or policy preferences could play a causal role, arguing they are epiphenomenal of partisanship (see Fowler 2020 for review). However, although voters are more likely to attitudinally sort by changing their views to align with party than vice versa (Carsey and Layman 2006; Dyck and Pearson-Merkowitz 2023; Levendusky 2009), both processes are consistent with sorting causing polarization, because both result in partisans having fewer attitudes in common with their opponents—the key mechanism of belief congruence theory.

MEASURING SORTING

Prior research has not placed these accounts in competition to assess how much they can explain rising affective polarization. This is likely due to the lack of comparable measures of demographic and attitudinal sorting. This paper overcomes that limitation by building on Brader et al.'s (2014) approach to measuring cross-pressures: Compile all relevant variables and gauge how well they predict partisanship (see also Harteveld 2021; Guntermann 2022). Well-sorted partisans will be confidently classified, whereas poorly sorted partisans will be classified with less confidence or inaccurately.

Using prediction models ensures sorting measures are comparable across time *and* domain. For example, I calculate how well individuals' partisanship can be predicted based on demographics in 1980 and 2024 *and* based on policy preferences in 1980 and 2024. These results are directly comparable because, no matter the year or variables involved, the model answers the same question: "How well do these traits predict partisanship?" Unlike additive formulas, predictive models also weight variables automatically. For example, they can account for religion having a stronger relationship with party ID than gender, as well as African Americans skewing more heavily Democratic than whites skew Republican. And by estimating separate models for each year, I account for coalitional changes over time without exercising discretion—the models automatically capture shifts ranging from the South's embrace of the Republican Party to the increase in non-whites' share of the electorate.

For these predictions, I follow Guntermann (2022) and apply random forest models—a machine-learning algorithm that classifies subjects using a collection of independent decision trees[5] (Montgomery and Olivella 2018). Random forests outperform traditional regressions because they automatically account for variable interactions and nonlinear relationships. I take the difference between respondents' estimated probabilities of belonging to their own party and their out-party.[6] Rescaling the result from 0–1 yields a continuous sorting measure, with higher scores indicating the respondent has more in common with their own party than the opposing party. For instance, a Republican whose demographics make him most likely to be a Republican gets a demographic sorting score of 1.0, whereas a Republican whose affiliations make her maximally likely to be a Democrat receives a 0.0. Appendix A demonstrates the stability and validity of these scores (pp. A.1–A.2).

[5]I use the ranger R package. All models generate 250 trees that predict respondents' partisanship using random subsets of data and predictor variables. I re-weight training data to equally represent Democrats and Republicans, which ensures sorting trends are not artifacts of shifting partisan balances. To avoid overfitting, I use five-fold cross-validation (every model uses five forests, each trained on 80% of the sample to estimate probabilities for the remaining 20%). The number of predictors per branch and minimum node size are jointly calibrated for each fold. Finally, I minimize measurement error by running five of these models for each sorting estimate, averaging each respondent's five estimates to calculate final sorting scores.

[6]I classify partisan leaners as members of their party, given their similarity to open partisans (Keith et al. 1992). I include pure independents as a third outcome in the model, but results do not differ when excluding them (Figure OA.4, p. A.4).

I address missingness by applying multivariate imputation by chained equations—a technique that estimates variables' empty values based on relationships with other variables (van Buuren and Groothuis-Oudshoorn 2011). Each year's imputations are run separately for demographic, ideological, policy, and group sentiment variables.[7]

Importantly, this sorting measure captures the overall ability of demographics to predict partisanship, not the intensity of specific group alignments (but see Figure 4). Thus, demographic sorting levels can remain stable amid considerable shifts in groups' partisan alignments, so long as the net predictive power of all demographics—the aggregate strength of all demographic cross-pressures—does not change.

Data Sources

I use the American National Election Studies (ANES) time series, which provides a consistent set of demographic measures dating back to 1952, the same gauges of ideology since 1972, and the same party feeling thermometers since 1980. Because feeling thermometer polarization levels differ significantly by survey mode (Figure 1; see also Tyler and Iyengar 2024; Thornton 2025), I exclude all online responses, including all of 2020. Appendix E analyzes the online time series (pp. A.5–A.6).

I calculate demographic sorting from 1952–2024 using all consistently available items (see Figure 2a). To ensure valid over-time comparisons, I maintain the same variable formats throughout. For instance, though the ANES has expanded its racial categories, I code all respondents who are not white, Black, or Hispanic as "other race" to match early years. For ideological sorting, I use seven-point ideological identification and feeling thermometer ratings of

[7] See Appendix D for details (pp. A.4–A.5).

liberals and conservatives.[8] Past research has used ideological feeling thermometers as a proxy for traditional ideology—combined, they correlate with ideology at approximately 0.6 (Bafumi and Shapiro 2009; Ellis and Stimson 2012; Wattenberg 2019).

Figure 2: Items used to measure partisan sorting

(a) Predictors for demographic and ideological sorting measures

Demographic sorting (12 fixed items)			**Ideological sorting (3 fixed items)**
Gender	Religion	Census region	Symbolic ideology
Race	Church attendance	Marital status	Conservatives FT
Age	Education	Union membership	Liberals FT
Income	Urbanicity†	Immigrant family	

(b) Predictors for attitudinal sorting measures beyond ideology

Sorting by attitude object

Sorting by attitude domain

	Policy sorting (top 14 items each year)	**Group sentiment sorting (top 14 items each year)**
Economic attitude sorting (top 8 items each year)	Includes Government jobs program Government health insurance Minimum wage	Includes Big business FT Labor unions FT Poor people FT
Cultural attitude sorting (top 8 items each year)	Includes Abortion Death penalty Climate action	Includes Gays and lesbians FT Feminists FT Evangelicals FT
Racial attitude sorting (top 8 items each year)	Includes Aid to Blacks Immigration Affirmative action	Includes Blacks FT Illegal immigrants FT Civil rights leaders FT

Partisan sorting is measured by using these sets of items to predict party ID. (Figures A1–A5 list all attitudes chosen each year.)
†The 2024 ANES lacks the regular urbanicity item, so I measure it using congressional districts instead. This does not impact results (see Figure OA.3, p. A.3).

This measure of demographic sorting does not incorporate demographic identity strength. Identity strength is likely an important dimension of partisan sorting, because it accounts for the subjective pressures voters feel from their group affiliations. For instance, a Black voter who

[8]Respondents who say they don't know or "haven't thought much" about their ideology or other attitudes are coded at 0.5. Refusals are coded as missing.

places little importance on their racial identity should feel less drawn to the Democratic Party than someone with a strong racial identity. However, data limitations prevent us from tracking the relationship between partisanship and demographic identity strength over time. This paper therefore cannot investigate whether "demographic identity" sorting has increased, focusing instead on demographic sorting based on group affiliations. Nevertheless, Analysis #4 replicates prior work to find that demographic identity strength's relationship with affective polarization is far smaller than political identity strength's.

Sorting by Attitude Object versus Attitude Domain

The ANES has few policy or group-sentiment items used consistently from 1972–2024. Therefore, I construct all other forms of attitudinal sorting using the *most politically relevant* attitudes each year, holding the number of predictors constant to maximize comparability over time. Specifically, I use preliminary random forests to identify the items each year that best predict partisanship. These variables are then placed in the main predictive models that calculate sorting (see Figures A1–A5). This approach, though atheoretical, guards against bias by selecting items in a consistent, results-agnostic manner.

As Figure 2b shows, all policy items are potential predictors for the policy sorting measure, and all group feeling thermometers are assigned to the group sentiment models. These sets of items both cover a wide range of economic and social debates, but they are each defined by their distinct target of evaluation, or their *object*: Policy variables measure prescriptive preferences, whereas group sentiments capture affective evaluations of social groups.

In addition, each policy and group sentiment item is assigned to one of three substantive *domains* of political conflict: economics, culture, and race.[9] Thus, this paper provides two specifications of the same analyses by categorizing items both by attitude object and domain. Crucially, policy and group sentiment sorting measures are kept in separate regressions from economic, cultural, and racial attitude sorting, ensuring no redundant predictors undermine analyses.

I employ both categorizations because, although prior sorting research mainly focuses on either policy preferences or group sentiments, these are not fully distinct categories and often cover similar topics. For instance, 1972–1988 measures of group sentiment sorting include a feeling thermometer for civil rights leaders. More broadly, many policy preferences—from gay marriage to immigration—are based primarily on attitudes about social groups. Indeed, policy preferences are often a consequence of group sentiments (Gilens 1999; Newman et al. 2026; Ramirez and Peterson 2020; Tesler 2016; cf. Clifford et al. 2025). Given this, are policy sorting's effects better understood as additional, indirect effects of group sentiment sorting? Arguably, readers should not interpret "policy sorting" as fully distinct from "group sentiment sorting," but instead as a related construct measured through a different format—one that de-emphasizes personal affect but captures many of the same group-oriented divisions.

Ultimately, the influence of attitudinal sorting documented in this paper reflects, in substantial part, the role of group-inflected attitudes in structuring partisan divisions. Because categorizing attitudes by their object can obscure this, I also measure sorting by attitude domain. This tracks divisions around substantive debates, regardless of whether each attitude is expressed

[9]Foreign policy items are not assigned to any domain. Following O'Brian (2024), I categorize environmental attitudes as cultural attitudes.

as a policy preference or group sentiment. As we will see, tracking the predictive power of cultural and racial attitudes demonstrates group-inflected attitudes' increasing importance.[10]

Finally, demographic sorting's variables are fixed, but attitudinal sorting's items vary over time. This asymmetry presents two possible concerns: I may underestimate changes in demographic sorting by excluding emerging groups, or I may show changes in attitudinal sorting that result from improved item coverage instead of true change.[11] However, these limitations do not drive my results. Expanding demographic predictors to include newly salient traits (e.g., sexual orientation) only increases demographic sorting levels by 0.01–0.02 points (Appendix B, pp. A.2–A.3), and attitudinal sorting measures constructed from fixed item sets produce similar trends from 1992–2024 (Appendix J, pp. A.14–A.16).

ANALYSIS #1: HAS PARTISAN SORTING INCREASED?

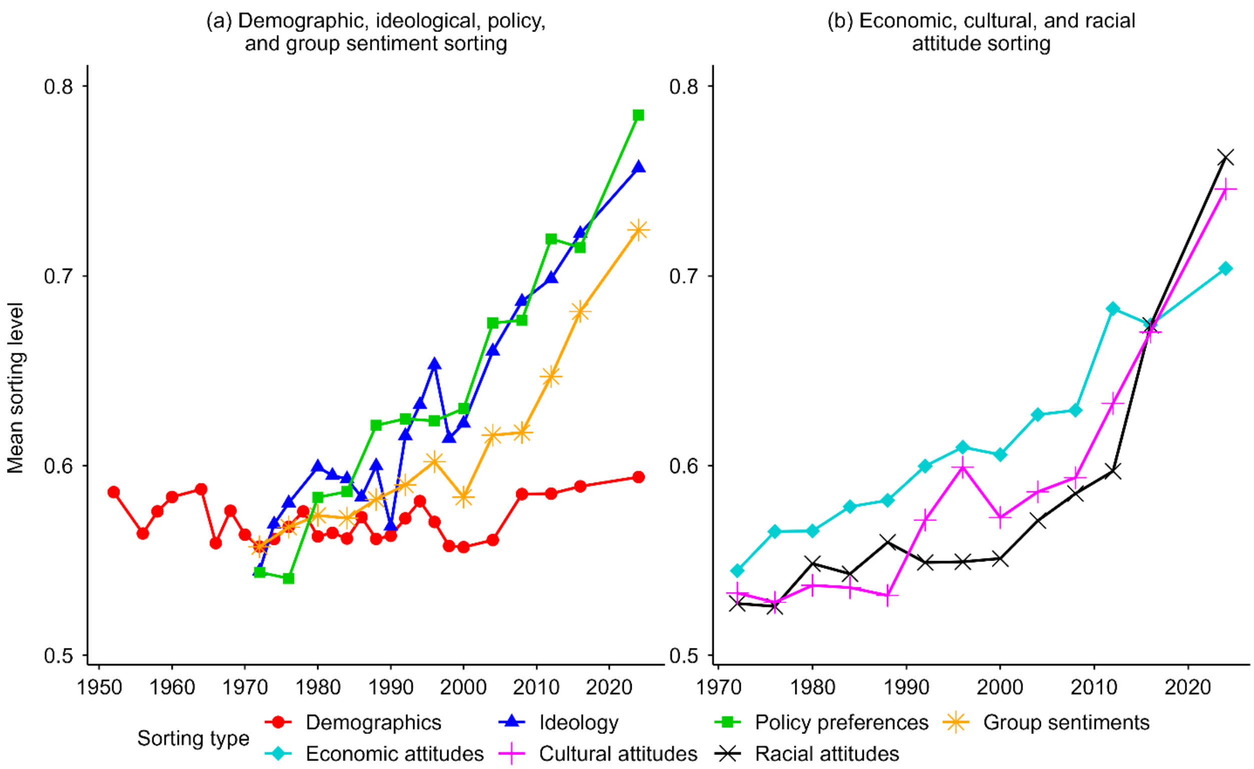
Figure 3 tracks mean sorting levels over time, and Table 2 presents comparisons across years. I find that demographic sorting has not meaningfully changed over time. At most, we see a modest uptick since 2004, but contemporary sorting levels remain indistinguishable from the mid-twentieth century. By contrast, all forms of attitudinal sorting have increased substantially—by at least four times as much as demographic sorting since 1972.

In the 1970s, voters were less sorted by policy preferences than by demographics, but now policy views are the most predictive of party ID. This suggests early scholars of public opinion, such as Campbell et al. (1960), were correct when they emphasized the importance of group affiliations over policy debates, but circumstances have changed considerably. Similarly,

[10]However, even this approach cannot account for group sentiments' influence across domains, such as racial views influencing economic attitudes (Gilens 1999; Tesler 2016).
[11]For instance, due to data limitations, economic attitude sorting measures earlier in the time series rely disproportionately on feelings toward economic classes (e.g., "poor people", "middle class people"), before the ANES introduced more economic policy items in the 1980s (see Figure A3).

economic attitude sorting has increased steadily over time—and was the most predictive domain for decades—but cultural and racial attitude sorting have both spiked since 2008. As of 2024, voters are now more sorted along racial and cultural lines compared to economics.

Figure 3: Partisan sorting levels over time



Mean partisan sorting scores in the ANES (face-to-face interviews) each year. Trends are split into two plots to improve readability. Items used to calculate sorting are listed in Figure 2 and Figures A1–A5.

Appendix F (pp. A.7–A.10) investigates whether rising levels of survey nonresponse could inflate attitudinal sorting estimates (Bailey 2024; Cavari and Freedman 2018, 2023; cf. Mellon and Prosser 2021, 2025). I demonstrate that this is a real concern but is, at a minimum, quite unlikely to account for a majority of attitudinal sorting's rise.

Overall, demographics' ability to predict party ID has been remarkably stable since 1952. It logically follows that the parties have not become more demographically divided over this period, so the rise of affective polarization cannot be explained by a corresponding rise in demographic sorting.

Table 2: Are over-time increases in sorting significant?

Sorting type	1952–2024	1972–2024	1980–2024
Demographic	0.01 (−0.01, 0.02)	0.04 (0.02, 0.05)	0.03 (0.02, 0.05)
Ideological	–	0.21 (0.20, 0.23)	0.16 (0.14, 0.18)
Policy	–	0.24 (0.23, 0.26)	0.20 (0.19, 0.22)
Group sentiment	–	0.17 (0.15, 0.18)	0.15 (0.13, 0.17)
Economic attitude	–	0.16 (0.15, 0.17)	0.14 (0.12, 0.15)
Cultural attitude	–	0.21 (0.20, 0.23)	0.21 (0.19, 0.22)
Racial attitude	–	0.24 (0.22, 0.25)	0.21 (0.20, 0.23)

Over-time changes in mean sorting levels in the ANES (face-to-face interviews). Parentheses contain 95% confidence intervals from weighted two-sided t-tests.

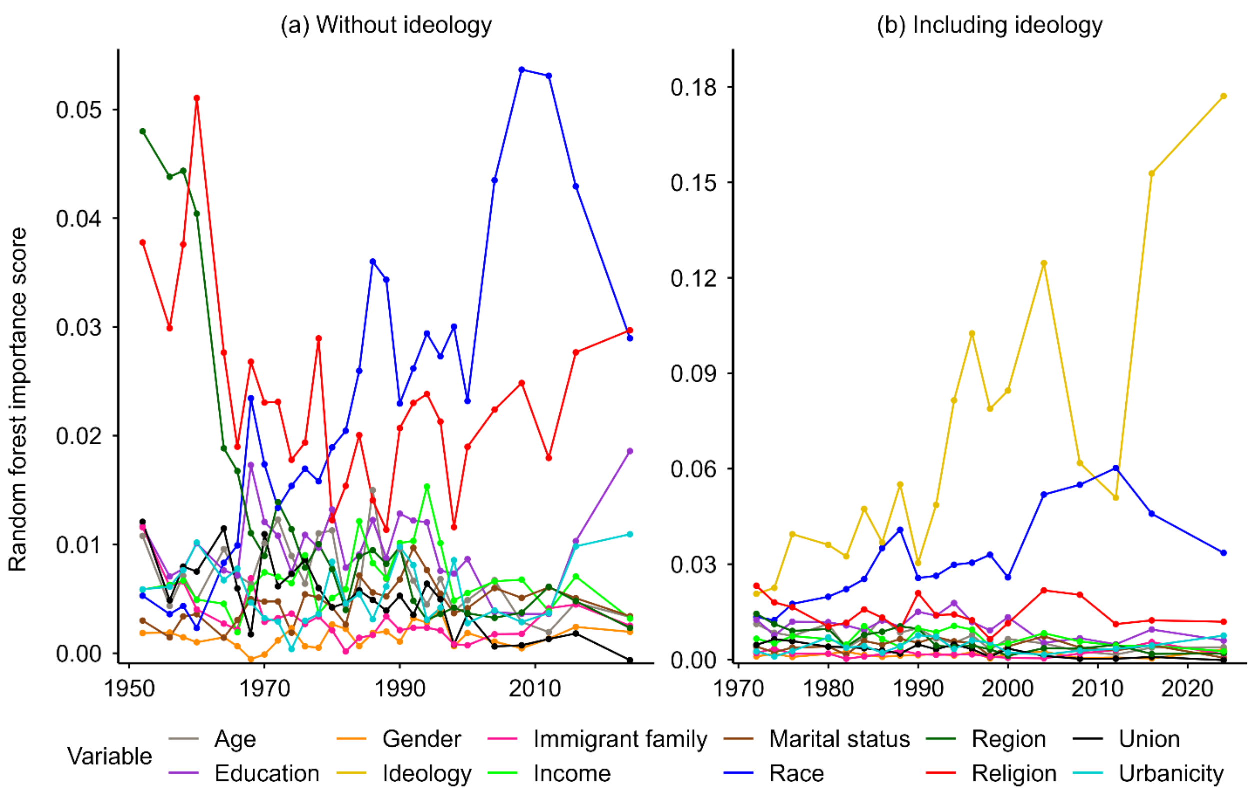
Which Demographics Are Most Sorted, and When?

Although demographic sorting has not increased overall, certain demographics have become more or less predictive of partisanship over time. I capture these coalitional shifts by calculating random forest feature importance scores[12] each year (see Figure 4a). Race has become a stronger predictor due to non-whites' increasing share of the electorate and disproportionate support for Democrats. However, this is offset by the end of the one-party South, which significantly reduced regional divisions. Furthermore, race's predictive power peaked in 2012 and has since declined substantially, which is consistent with claims of racial depolarization (Chotiner 2024; Ruffini 2023). Religion's significance declined as the divide between Protestants and Catholics faded but has since resurged because of a divide between churchgoers and non-

[12]Feature importance scores measure the loss of predictive power when scrambling a variable's values—greater losses signify higher importance.

churchgoers. Education and urbanicity have, unsurprisingly, become the third and fourth most predictive variables.

Figure 4: Social group importance over time



The abilities of various social groups to predict party ID (as measured by random forest importance scores), over time. Religion combines religious ID and church attendance scores. All scores in panel (b) come from separate random forests that combine demographics and three-point ideology.

These results contextualize current sociopartisan divisions. There are significant demographic differences—and clear ongoing shifts—between the parties. But the magnitude of these divides is not new: Individual groups become more or less pivotal, and some traits even predict opposite party IDs at different times. For instance, low education and Southern residence each strongly predicted being a Democrat in the 1950s but now predict Republican identification (Table OA.1, p. A.2). Thus, emerging partisan alignments have merely replaced old ones, and the parties have remained similarly heterogeneous throughout this demographic churn.[13]

[13]See Lipset (1960, pp. 285–309) on partisan-demographic divides in the mid-twentieth century and earlier.

Furthermore, Figure 4b shows ideology has become far more effective in predicting partisanship than any demographic. This is notable, given the tendency of prior sorting research to treat ideology as one of many aggregated social groups (Jin et al. 2025; Mason 2018; Mason and Wronski 2018). This practice may overstate demographic sorting's importance. This paper is agnostic on whether symbolic ideology is more of a social group or a meaningful worldview, but future research should, at a minimum, distinguish between ideological and demographic sorting.

ANALYSIS #2: HAS SORTING INCREASINGLY CAUSED POLARIZATION?

Even if demographic sorting has not increased, it could still drive rising affective polarization if sorting's impact on polarization has intensified over time—a serious possibility in an era of "identity-inflected" politics (Sides et al. 2018). This claim is difficult to definitively test, but observational evidence suggests it is unlikely.

I regress affective polarization (the difference between in-party and out-party feeling thermometer ratings, rescaled from 0–1) on each form of sorting, using the following specifications:

$$\mathrm{Polarization_i} = \alpha_0 + \textstyle\sum_{\mathrm{d}} \alpha_{\mathrm{d}} \mathrm{Decade_d} + \alpha_{\mathrm{s}} \mathrm{Sorting_i} + \varepsilon_{\mathrm{i}} \quad (1)$$

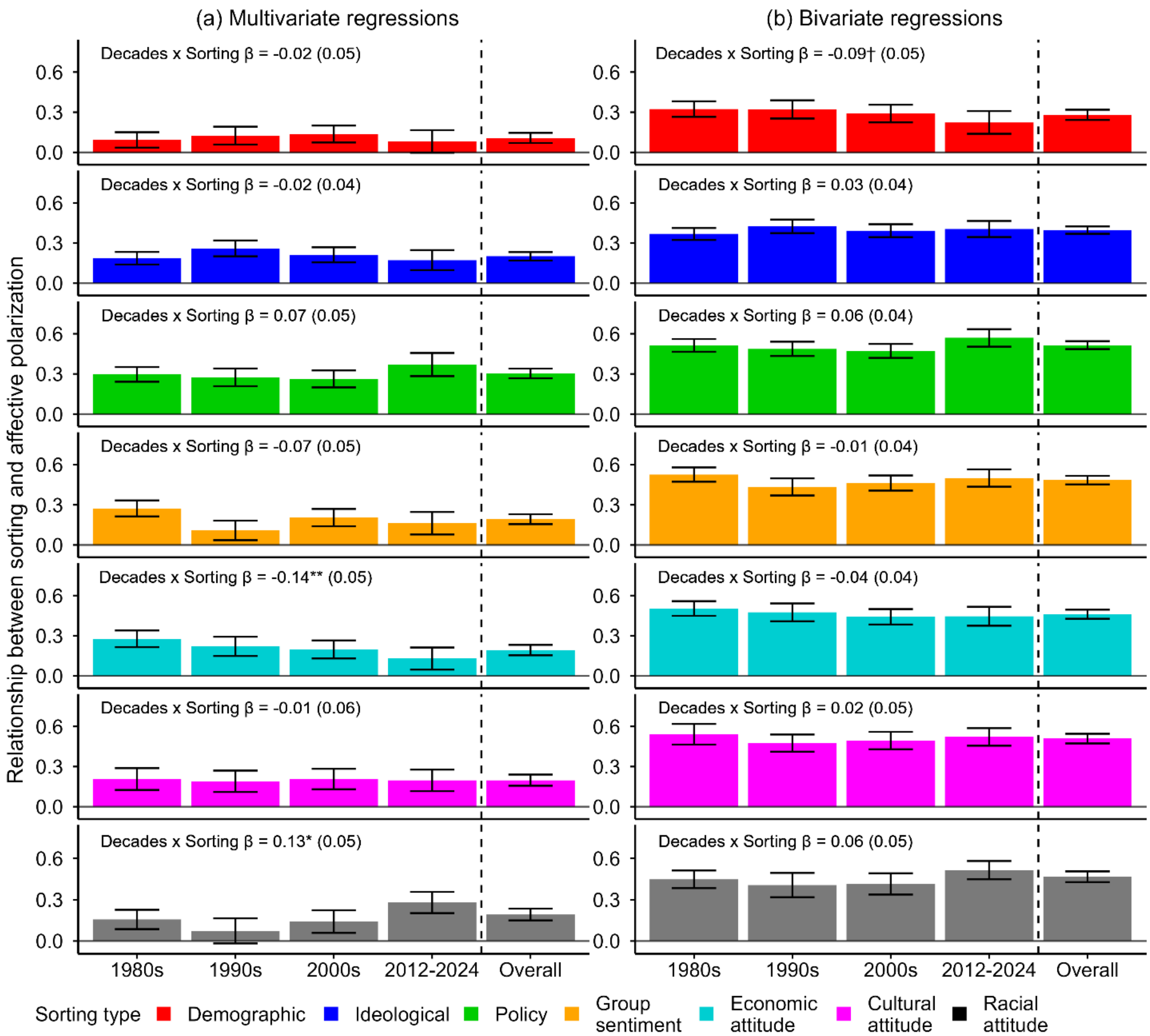
$$\mathrm{Polarization_i} = \delta_0 + \textstyle\sum_{\mathrm{d}} \delta_{\mathrm{d}} \mathrm{Decade_d} + \beta_{\mathrm{s}} \mathrm{Sorting_i} + \sum_{\mathrm{d}} \gamma_{\mathrm{d}} (\mathrm{Decade_d} \times \mathrm{Sorting_i}) + \varepsilon_{\mathrm{i}} \quad (2)$$

$$\mathrm{Polarization_i} = \beta_0 + \beta_1 \mathrm{Decades} + \beta_2 \mathrm{Sorting_i} + \beta_3 (\mathrm{Decades} \times \mathrm{Sorting_i}) + \varepsilon_{\mathrm{i}} \quad (3)$$

$\mathrm{Decade_d}$ represents dummy variables for the 1990s, 2000s, and 2012–2024. Equation #1 captures the overall relationship between sorting and polarization (α_s). Equation #2 allows the relationship to vary by decade, with β_s representing the baseline 1980s relationship and each γ_d capturing changes in subsequent decades. Equation #3 interacts sorting with a linear Decades variable (β_3)

scaled from 0 (1980s) to 1 (2012–2024), which tests whether sorting's relationship with polarization has significantly changed over time.[14]

Figure 5: Partisan sorting's relationship with affective polarization over time



Bars show regression coefficients from Equations #1 ("Overall") and #2 (each decade). Decades × Sorting β presents Equation #3's interaction coefficients. Affective polarization is the dependent variable in all models. Bivariate models use a single form of sorting, and Multivariate models use Demographic and Ideological sorting alongside either Policy and Group Sentiment sorting or Economic/Cultural/Racial Attitude sorting. (Source: ANES, face-to-face interviews.) $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

For each type of sorting, I run both a "Bivariate" model that uses only that form of sorting and a "Multivariate" version that includes demographic and ideological sorting alongside all

[14]Time variables are binned by decade to improve power. Figure OA.12 bins by year (p. A.13).

measures of sorting by either attitude object or attitude domain.[15] Although these regressions use no other controls, Appendix I discusses the inclusion of demographic and attitudinal predictors as controls to account for their "direct effects" on polarization, and it demonstrates the significant but modest impact of controlling for partisanship, party ID strength, political knowledge, and political interest (pp. A.13–A.14). Finally, I upweight years with smaller samples and downweight years with larger samples, ensuring all years contribute equally to the results.

Figure 5 displays the results. The relationships are largely stable, though racial attitude sorting has likely become more predictive of polarization, and economic attitude sorting may have become less predictive. Critically, neither specification suggests demographic sorting's relationship with polarization has strengthened. Thus, the second necessary condition—a strengthening relationship between demographic sorting and polarization—is also unmet.

Considering Demographic Sorting's Potential Indirect Influence

Demographic sorting could still indirectly drive rising affective polarization by causing attitudinal sorting to increase, which in turn increases polarization. Group affiliations can strongly influence policy preferences and group sentiments (Conover 1988; Kinder and Sanders 1996), and this influence typically occurs early in life—earlier in Campbell et al.'s "funnel of causality" than policy preferences or even partisanship (1960, pp. 24–37). Given this, cross-sectional analyses could underestimate demographic sorting's importance by attributing its effects to more proximate factors.

I address this concern in three ways. First, Figure 5's Bivariate regressions capture both direct and indirect influence for each form of sorting, including demographic sorting's influences

[15]In Analyses #2–3, "Multivariate" results for demographic and ideological sorting are based on regressions with sorting by attitude object. Results do not change when using attitude domains.

through attitudes. Although this yields a stronger overall relationship with polarization, the relationship marginally weakens over time, rather than strengthening.

Figure 6: The over-time correlations of partisan sorting

Demographic sorting's yearly relationships with each form of attitudinal sorting (panel a), and the strength of each form of attitudinal sorting with all others, by decade (panel b). Cells contain weighted Pearson correlations. (Source: ANES, face-to-face interviews.)

Second, if demographic sorting has increasingly caused attitudinal sorting, they should become more strongly correlated over time. If anything, Figure 6a shows the opposite: Demographic sorting's correlations with attitudinal sorting are typically weak and, in some cases, appear to grow weaker over time. By contrast, Figure 6b shows stronger relationships between forms of attitudinal sorting, some of which strengthen over time. Thus, demographic sorting does not appear to have indirectly caused rising polarization by increasing attitudinal sorting.

Third, Appendix G uses the Youth-Parent Socialization Study to test whether early-life demographic sorting predicts later affective polarization and attitudinal sorting (pp. A.10–A.12). I find modest short-run associations, but few persist in the long run. The results may inform future research on political socialization and the initial development of polarized attitudes (see also Tyler and Iyengar 2023), but they provide little evidence that early-life demographic sorting has contributed meaningfully to the rise of affective polarization.

ANALYSIS #3: ESTIMATING PARTISAN SORTING'S MAXIMUM POTENTIAL CONTRIBUTIONS

Because all forms of partisan sorting are measured comparably, I can now quantify how much each form *could have* contributed to rising polarization—assuming the relationship between sorting and affective polarization is fully and unidirectionally causal. This is an unrealistic assumption (see Analysis #5), but it allows us to properly contextualize the negligible changes in demographic sorting and to explore the relative importance of different attitudes.

I estimate the increase in affective polarization predicted by the change in sorting levels from 1980–2024 shown in Analysis #1 *and* the strength of relationship[16] between sorting and polarization shown in Analysis #2. I employ three regression models:

$$\mathrm{Polarization_i} = \delta_0 + \delta_1 \mathrm{Time_i} + \varepsilon_i \quad (4)$$
$$\mathrm{Sorting_i} = \gamma_0 + \gamma_1 \mathrm{Time_i} + \varepsilon_i \quad (5)$$
$$\mathrm{Polarization_i} = \beta_0 + \beta_1 \mathrm{Sorting_i} + \mathrm{Controls_i} + \varepsilon_i \quad (6)$$

Time is rescaled from 0 (1980) to 1 (2024). The coefficient δ_1 estimates the increase in affective polarization over time, γ_1 estimates the change in sorting levels, and β_1 estimates the relationship between sorting and polarization. The Multivariate model includes other forms of partisan sorting as controls (as in Analysis #2), and the Bivariate model includes no controls. With these models, I estimate each type of sorting's contribution to increased polarization:

[16]The ANES lacks sufficient power to precisely estimate interaction effects, and adding these imprecise estimates into the decompositions expands confidence intervals beyond usefulness. Therefore, I assume a stable relationship in Analysis #3—an assumption largely consistent with prior results.

$$\text{Sorting's Contribution \%} = \left(\frac{\gamma_1 \times \beta_1}{\delta_1}\right) \times 100$$

For example, the Multivariate demographic sorting regressions yield:

$$\text{Polarization} = 0.253 + 0.164 \times \text{Time} \qquad (\delta_1 = 0.164)$$
$$\text{Sorting} = 0.557 + 0.034 \times \text{Time} \qquad (\gamma_1 = 0.034)$$
$$\text{Polarization} = -0.193 + 0.106 \times \text{Sorting} + \text{Controls} \qquad (\beta_1 = 0.106)$$

With these values, I estimate that demographic sorting can plausibly explain 2.20% of the rise in affective polarization since 1980. By contrast, each form of attitudinal sorting can explain between 16.56% and 36.42% of the increase in affective polarization (Table 3). Even when using Bivariate regressions to account for indirect effects, I estimate that demographic sorting can only explain 6.14%.

Appendix I replicates these calculations across several specifications, including different regression controls in Equation #6 (pp. A.13–A.14). While multivariate model estimates change modestly across specifications—particularly when controlling for political knowledge and interest—the inclusion of additional controls reduces the demographic sorting Bivariate estimate by nearly half (from 6.14% to 3.46%) and all Bivariate attitudinal estimates by 15–20 percentage points.

To be clear, this analysis serves only as a thought experiment to contextualize Analyses #1–2. The "Maximum Contribution" estimates represent unrealistic upper bounds, not identified causal estimates—even the confidence intervals represent uncertainty within the context of this thought experiment and cannot account for alternate causal pathways. Analyses #5a–5b present evidence of causality in *both* directions, and there is no way to credibly adjust Table 3's estimates to reflect this bidirectionality or additional confounding factors (such as elite cues' joint effects on sorting and polarization).

Table 3: How much of affective polarization's increase can sorting explain? (1980–2024)

	Multivariate regressions			Bivariate regressions	
Sorting	**γ_1**	**β_1**	**Maximum contribution**	**β_1**	**Maximum contribution**
Demographic	0.03 (0.01)	0.11 (0.02)	2.20% (1.06% – 3.34%)	0.30 (0.02)	6.14% (3.70% – 8.58%)
Ideological	0.17 (0.01)	0.20 (0.02)	21.03% (17.21% – 24.84%)	0.42 (0.01)	43.23% (38.29% – 48.18%)
Policy	0.19 (0.01)	0.31 (0.02)	36.42% (31.44% – 41.40%)	0.54 (0.01)	62.48% (56.76% – 68.21%)
Group sentiment	0.14 (0.01)	0.19 (0.02)	16.81% (13.18% – 20.44%)	0.51 (0.02)	44.25% (39.05% – 49.45%)
Economic attitude	0.14 (0.01)	0.19 (0.02)	16.56% (12.90% – 20.22%)	0.49 (0.02)	41.85% (36.90% – 46.80%)
Cultural attitude	0.19 (0.01)	0.20 (0.02)	22.64% (17.70% – 27.57%)	0.54 (0.02)	61.28% (55.50% – 67.06%)
Racial attitude	0.18 (0.01)	0.19 (0.02)	21.18% (16.26% – 26.10%)	0.50 (0.02)	54.26% (48.30% – 60.22%)
All attitudes (by object)			**74.25%** **(67.00% – 81.51%)**		
All attitudes (by domain)			**81.40%** **(72.65% – 90.15%)**		

Back-of-the-envelope estimates of how much each form of partisan sorting could have contributed to rising affective polarization since 1980. Estimates use individual-level ANES data (face-to-face interviews). γ_1 represents sorting's over-time change. β_1 represents sorting's overall relationship with affective polarization. "All attitudes (by object)" combines ideological, policy, and group sentiment sorting; "All attitudes (by domain)" combines ideological, economic, cultural, and racial attitude sorting. 95% confidence intervals use the delta method to account for uncertainty in both γ_1 and β_1. (Source: ANES, face-to-face interviews.)

Furthermore, readers should compare attitudinal sorting's estimated contributions with caution, given differences in measurement quality. For instance, social desirability bias may lead respondents to answer certain feeling thermometers less honestly, muting partisan differences in group sentiments. Conversely, if voters increasingly regard "liberal" and "conservative" as synonyms for "Democrat" and "Republican," this could arbitrarily inflate symbolic ideological divides.

Nevertheless, the contrast between demographic and attitudinal sorting's plausible contributions is stark. Observational data provide little evidence that demographic sorting has contributed to the rise in affective polarization. This does not mean demographic sorting has zero impact on current polarization levels—social group memberships clearly do matter. Consistent with Campbell et al. (1960), Converse (1964), Achen and Bartels (2017), and Mason (2018), the parties have somewhat distinct demographic profiles, these profiles stably influence affective polarization levels, and demographics likely exert indirect influence through their impacts on attitudes. But these dynamics cannot explain an *increase* in polarization.

ANALYSIS #4: REVISITING PRIOR WORK ON SOCIAL SORTING

These conclusions differ from conventional wisdom (see Table 1)—most notably Mason's (2018) landmark *Uncivil Agreement*. However, this work focuses on the combined influence of demographic group memberships, demographic identity strength, ideological identity, and partisan identity under the broad umbrella of "sociopartisan sorting." Thus, Mason (2018) does not claim demographic sorting is the *only* source of rising polarization, attributing it instead to strengthened and politicized identities, broadly conceived. Within this framework, demographics have a central but non-exclusive role: "As the parties became more homogeneous in ideology, race, class, geography, and religion, partisans on both sides felt increasingly connected to the groups that divided them" (Mason 2018, p. 40).

However, this framework does not differentiate between objective demographic sorting, subjective demographic sorting, ideological sorting, and partisan identity. Instead, the book's key analyses combine six social and political affiliations into an overall additive index, with each affiliation weighted by identity strength (see Table OA.10, p. A.17). Specifically, the index combines partisan, ideological, evangelical, secular, African American, and Tea Party identities—

one racial identity, two religious identities, and three political identities. The book's analyses mostly control for the scale's partisan identity component, but they do not separate the other political identities' contributions. Thus, although this measure captures the broad phenomenon of politicized identities, it may not reflect *demographic* sorting.

Table 4: Sociopartisan sorting and affective polarization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sociopartisan sorting (scale)	0.36**	0.65**								
	(0.08)	(0.08)								
Political identities										
Political identity sorting (subscale)			0.70**				0.80**			
			(0.12)				(0.11)			
Partisan identity strength	0.31**			0.39**				0.44**		
	(0.05)			(0.05)				(0.04)		
Ideological identity sorting				0.08†				0.09*		
				(0.05)				(0.04)		
Tea Party identity sorting				0.08				0.09		
				(0.09)				(0.09)		
Demographic identities										
Demographic identity sorting (subscale)			0.10†		0.14**		0.11*		0.18**	
			(0.05)		(0.05)		(0.04)		(0.04)	
Evangelical identity sorting				0.03		0.10		0.04		0.16**
				(0.07)		(0.06)		(0.06)		(0.06)
Secular identity sorting				0.07		0.09		0.09†		0.15**
				(0.06)		(0.06)		(0.05)		(0.05)
Black identity sorting				0.18†		0.26**		0.15		0.24*
				(0.11)		(0.10)		(0.11)		(0.11)
Issue extremity	0.21**	0.22**	0.23*	0.19*	0.31**	0.28**				
	(0.07)	(0.08)	(0.10)	(0.08)	(0.11)	(0.08)				
Issue constraint	−0.01	−0.02	0.02	0.03	0.05	0.07				
	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)	(0.05)				
Intercept	−0.19**	−0.14†	−0.30**	−0.21*	0.04	−0.10	−0.20†	−0.14	0.23*	−0.01
	(0.07)	(0.08)	(0.12)	(0.10)	(0.10)	(0.10)	(0.11)	(0.09)	(0.09)	(0.09)
Observations	775	697	424	729	424	745	503	877	503	896

This table assesses affective polarization's relationship with political and demographic identity sorting, based on the 2011 YouGov data used in Mason (2018); column #1 reproduces Table A.5. Omitted controls match the original analysis: political sophistication, education, race, gender, income, and church attendance. See Appendix K (pp. A.17–A.18) for details. $^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$

Table 4 reproduces and expands Mason's (2018) analysis. Columns #1–2 confirm the aggregate sociopartisan sorting index strongly predicts affective polarization, even when controlling for partisan identity strength. However, when dividing the index into demographic and political identity subscales or fully disaggregating it, political identities predict polarization far more strongly— demographic identities have only marginally significant predictive power unless political identities or policy attitudes are removed from the models.

Table 5: Sociopartisan sorting and affective polarization (additional replications)

	CCES module (2016)			ANES (2012)			Nationscape (2019–2021)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sociopartisan sorting (full scale)	0.63** (0.17)			0.80** (0.03)			0.63** (0.02)		
Political identities									
Political identity sorting (subscale)		0.58** (0.17)			0.58** (0.02)			0.53** (0.02)	
Partisan identity strength			0.17** (0.06)			0.27** (0.01)			0.17** (0.01)
Ideological identity sorting			0.27† (0.15)			0.35** (0.04)			0.44** (0.02)
Demographic identities									
Demographic identity sorting (subscale)		0.19† (0.11)			0.14** (0.03)			0.15** (0.02)	
Racial identity sorting			0.09 (0.08)			0.12** (0.02)			0.07** (0.01)
Religious identity sorting			0.12 (0.09)			0.03 (0.02)			0.06** (0.01)
Intercept	−0.05 (0.14)	−0.13 (0.15)	−0.00 (0.13)	−0.03 (0.03)	0.01 (0.03)	−0.02 (0.03)	−0.02 (0.02)	−0.04* (0.02)	−0.09** (0.02)
Observations	684	653	679	4,743	4,736	4,616	30,747	30,712	30,484

*This table assesses affective polarization's relationship with political and demographic identity sorting, based on the 2016 CCES module used by Mason and Wronski (2018), the 2012 ANES, and UCLA Nationscape data. Omitted controls match Table A.3 from Mason and Wronski (2018): gender, age, education, income, partisanship, and political interest and knowledge (ANES models also control for survey mode; Nationscape models include wave fixed effects). See Appendix K (pp. A.17–A.18) for details. †p < .10; *p < .05; **p < .01*

Table 5 extends this replication to Mason and Wronski (2018), who analyze the 2016 CCES University of Mississippi module using a similar sociopartisan sorting measure based on racial, religious, ideological, and partisan identities. Once again, political identity sorting has far stronger predictive power than demographic identity sorting (columns #1–3).[17] I then replicate these analyses using two larger samples: the 2012 ANES and seven UCLA Nationscape survey waves.[18]

[17]Mason and Wronski (2018) focus on in-party warmth rather than affective polarization; demographic identity sorting does not significantly predict in-party warmth.

[18]Both datasets include the variables needed to replicate Mason and Wronski's (2018) sociopartisan sorting measure: single-item measures of partisan, ideological, racial, and religious identity strength. The ANES analyses use folded party ID and ideology to measure identity strength, consistent with Mason (2015) and Mason and Wronski (2018). The Nationscape analyses also use folded ideology.

The same pattern holds: Demographic identity sorting is predictive of affective polarization, but far less strongly than partisan and ideological identity. Thus, although we do not know whether demographic identity sorting has increased over time, it is notable that prior work on social sorting and polarization provides stronger evidence for the importance of attitudes than for demographics, even when incorporating identity strength.

ANALYSIS #5A: INVESTIGATING CAUSALITY (PANEL DATA)

The above analyses suggest rising partisan differences in "who we are" have not caused rising polarization, but rising differences in "what we think" may have. However, observational analyses cannot show attitudinal sorting *causes* affective polarization—merely that it has risen over time and is associated with polarization.

Attitudinal sorting is often a product of voters changing their attitudes based on elite cues, rather than changing their party to better fit their attitudes (Carsey and Layman 2006; Dyck and Pearson-Merkowitz 2023; Levendusky 2009). When this conformity occurs, perhaps voters are adopting their party's positions, losing cross-cutting ties to the opposite party, and therefore becoming more hostile. However, partisan animus may also motivate those attitude changes, thereby causing increased sorting. Using original panel analyses and a meta-analysis of prior experiments, I assess these possibilities.

Prior panel analyses suggest both processes occur. Lelkes (2018) finds "a reciprocal, albeit weak, relationship" between policy sorting and affective polarization. Robison and Moskowitz (2019) find a strong causal role for group sentiment sorting in driving increased polarization and a weaker effect in the opposite direction. I conduct similar analyses on the 1972–1976, 1992–1996, and 2016–2024 ANES panels, as well as the Youth-Parent Socialization Study's 1973–1997 waves. For each panel, I use all relevant, consistently available items to calculate sorting scores for each wave (see Figure OA.16, p. A.18).

I use the following cross-lagged panel models (CLPMs) to estimate the causal effects of attitudinal sorting and affective polarization on each other (Finkel 1995):

$$\text{Polarization}_{i,t} = \beta_0 + \beta_1 \text{Sorting}_{i,t-1} + \beta_2 \text{Polarization}_{i,t-1} + \text{Controls}_{i,t-1} + \varepsilon_i \tag{7}$$

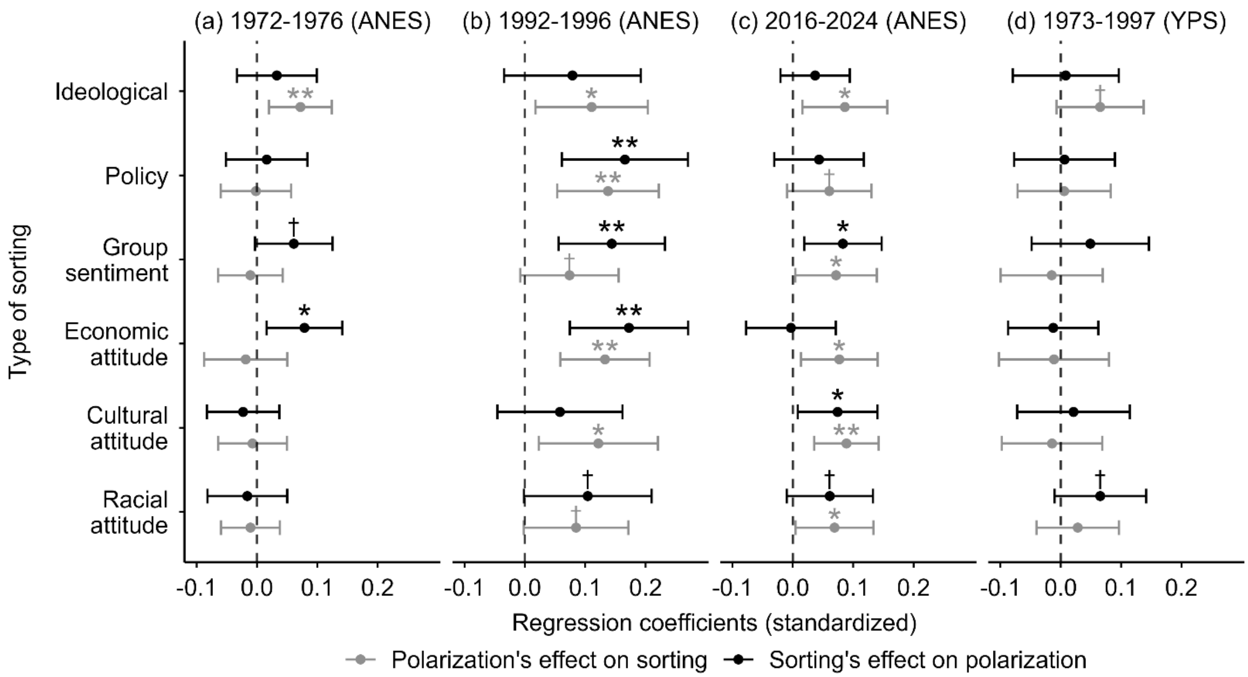
$$\text{Sorting}_{i,t} = \gamma_0 + \gamma_1 \text{Sorting}_{i,t-1} + \gamma_2 \text{Polarization}_{i,t-1} + \text{Controls}_{i,t-1} + \varepsilon_i \tag{8}$$

The results are broadly consistent with prior findings (Figure 7). Within the ANES panels, attitudinal sorting and affective polarization each show positive and at least marginally significant effects in at least half of all tests. Group sentiment sorting has the most consistent effect on polarization, whereas ideological sorting appears to be largely a consequence of polarization. The influence of economic attitude sorting appears to have faded in recent years, while cultural and racial attitude sorting's effects have likely strengthened. Affective polarization shows minimal influence in earlier panels, but it has since developed a consistent impact on sorting.[19] Overall, these analyses suggest attitudinal sorting and affective polarization are reciprocally predictive of each other, with identity-inflected attitudes in particular developing a stronger causal role over time.

Figure OA.18 reports wave-specific CLPMs (p. A.20), which allow cross-lagged effects to vary across pairs of waves. These models often show stronger effects than the main text's pooled estimates, particularly for 2016–2024. In addition, the 1973–1997 panel has a sparse set of policy and feeling thermometer items available across all waves, likely resulting in noisier estimates of attitudinal sorting. Thus, Figure 7 may represent conservative estimates.

[19]However, the 1972–1976 and 1973–1997 panels measure affective polarization using ratings of "Republicans" and "Democrats" rather than the standard "Republican Party" and "Democratic Party" thermometers. Polarization's changing impact could therefore be due to measurement differences.

Figure 7: The reciprocal relationships between attitudinal sorting and affective polarization



*Coefficient estimates from cross-lagged panel models applied to three ANES panels and the Youth-Parent Socialization (YPS) study. Figure OA.16 (p. A.18) lists items used to measure sorting. Regressions include demographic controls (and wave 1 survey mode for 2016–2024). Figure OA.18 (p. A.20) separates results into waves 1–2 and 2–3. †p < .10; *p < .05; **p < .01*

However, CLPMs also have inherent limitations that can inflate estimated causal effects. Despite common use in political behavior research (see Carsey and Layman 2006; Highton and Kam 2011; Margolis 2018), they do not account for trait stability and thus conflate between-person and within-person effects (Hamaker et al. 2015; Lucas 2023). Appendix M discusses this issue, presents results from random-intercept cross-lagged panel models, and covers those models' limitations (pp. A.19–A.21). Ultimately, these analyses provide suggestive evidence of causal roles for attitudinal sorting and affective polarization. Analysis #5b corroborates sorting's role with stronger internal validity.

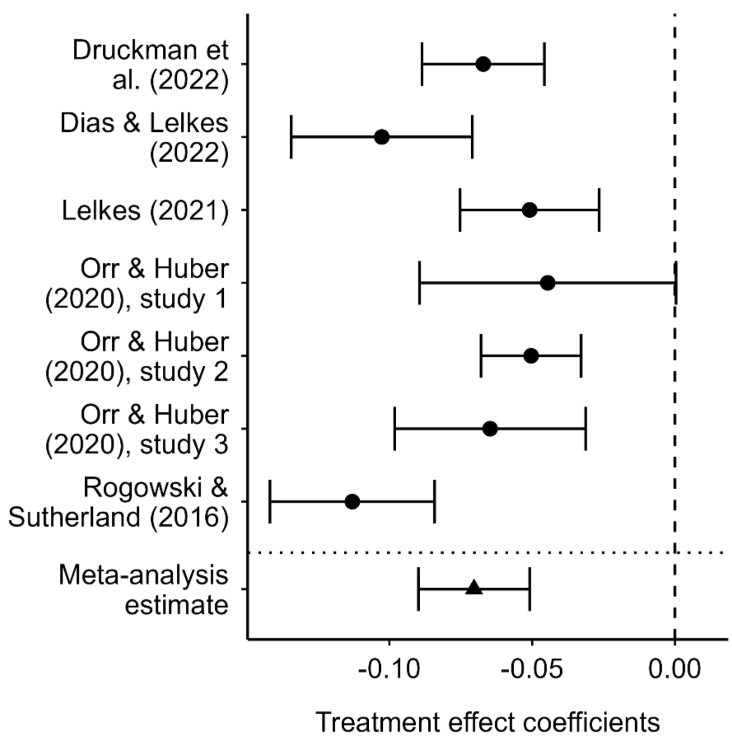
ANALYSIS #5B: INVESTIGATING CAUSALITY (META-ANALYSIS)

Finally, I conduct a meta-analysis of prior experiments that present respondents with hypothetical out-partisans, randomly assigning some with counter-stereotypical attitudes (Dias and

Lelkes 2022; Druckman et al. 2022; Lelkes 2021; Orr and Huber 2020; Rogowski and Sutherland 2016). For instance, one respondent might evaluate a hypothetical Republican with no policy information, and another rates a Republican who supports adoption for same-sex couples. Comparing these evaluations reveals how partisans would regard an out-party less defined by opposing policies.

I find counter-stereotypical views consistently reduce negative affect toward out-partisans (Figure 8). This indicates that voters' out-party animus is in part a consequence of the parties becoming more attitudinally sorted—partisans have fewer counter-stereotypical views that mitigate affective polarization. The result is consistent with research showing that raising the salience of policy disagreements raises out-party distrust (Gidron et al. 2025).[20]

Figure 8: Meta-analysis of past experiments on sorting and affective polarization



A meta-analysis of prior experiments assessing the impact of counter-stereotypical policy information on out-party animus. Table OA.12 describes studies (p. A.21). Meta-estimate comes from a random effects model using R's metafor package.

[20]The meta-analysis has no studies about group sentiments due to data limitations only; this should not be interpreted as challenging their causal role.

Few studies examine the reverse effect of affective polarization on sorting. Broockman et al. (2023) and Levendusky (2023) exogenously reduce polarization and measure the effect on partisan cue-taking, which should in turn spur partisan sorting. The results are mixed: Broockman et al. (2023) find two null effects, whereas Levendusky (2023) finds a modest treatment effect ($p < 0.1$) and concludes, “While ideological polarization has a large effect on affective polarization, the reverse causal process—from affective polarization to ideological divisions—is considerably weaker” (p. 129).

Overall, the causal evidence suggests partisan sorting and affective polarization are linked, with broader evidence for sorting’s effect on polarization. However, these findings are not definitive. The experiments have stronger internal validity than panel studies, but their indirect nature limits external validity. Furthermore, it may be harder to experimentally manipulate affective polarization in a way that connects with sorting than vice versa. For instance, a cooperative out-partisan in a trust game may reduce polarization without manipulating perceptions of partisan policy disagreements.

Thus, these causal claims—insofar as they can be investigated with the tools of social science—require further study beyond this paper. Nevertheless, attitudinal sorting has likely contributed to rising affective polarization, and vice versa. The same cannot be said for demographic sorting.

DISCUSSION

Concerned observers often take for granted that America’s political parties have sorted along demographic lines, leading to historically high levels of affective polarization (see Table 1). This paper suggests otherwise, showing that the conditions necessary for this theory are simply not met: Demographic sorting has not increased since 1952, and it has not become more predictive of polarization. By contrast, attitudinal sorting—based on symbolic ideology, policy preferences, and group sentiments—has increased substantially and strongly predicts affective polarization. Prior research emphasizing the importance of demographic sorting uses a sorting measure that includes partisan identity and ideology; reanalyses show the latter two variables drive most of the results. Finally, panel analyses and a meta-analysis of experiments suggest attitudinal sorting has some causal effect on polarization, and vice versa.

This paper is not without limitations. The analyses do not investigate peer-network sorting. Politically diverse social networks predict lower partisan animus (Baldassarri and de Jong 2025; see also Mutz 2006; Rawlings 2022), but the ANES does not track this over time. The ANES also lacks measures of affective polarization beyond party feeling thermometers. Other measures better capture interpersonal animus (Campos and Federico 2026; Druckman and Levendusky 2019), and future research should assess their relationships with both demographic and attitudinal sorting.

I have taken pains to differentiate between sorting based on demographic *categories* versus demographic *identities*, and future work should continue investigating the latter's role. However, prior research has often elided the distinctions between group memberships, demographic identity importance, and non-demographic identities. Going forward, I hope this paper spurs more conceptual clarity on this point.

This paper measures demographic sorting as an aggregation of group memberships, but individual demographic-party alignments have shifted in different ways (see Figure 4). Perhaps sorting on specific demographics contributes to polarization differently than the net effect of all groups. However, given the stable relationship shown in Analysis #2, any polarizing effect of increased sorting on one demographic would need to be offset by depolarizing effects from another demographic's increased sorting. This is a possibility, but one without obvious theoretical grounding.

Analysis #3 suggests a stronger role for policy sorting compared to group sentiment sorting. However, as discussed above, the two phenomena are tightly intertwined, and the latter is likely inhibited by social desirability bias. Given that most policies affect social groups in some manner, future research looking to distinguish between these attitudes faces a considerable challenge.

Some scholars question whether affective polarization is truly a negative force in American democracy. For instance, reducing affective polarization may not ameliorate more tangible problems such as anti-democratic attitudes (Broockman et al. 2023; Voelkel et al. 2023; cf. Voelkel et al. 2024). Indeed, this paper's findings bolster arguments against alarm over affective polarization, because they suggest rising polarization is a response to the emergence of genuine differences between the parties—among elites *and voters*—rather than irrational tribalism (see also Orr et al. 2023; Fowler et al. 2026). Partisan animus likely presents real challenges in its own right (Druckman et al. 2023), but the root problem appears to be widening substantive disagreements between the parties.

Nevertheless, Appendix O extends Analysis #2 to additional dependent variables focused on actual behavior, as well as party ID stability and strength (pp. A.22–A.24). Table 6 summarizes the results: Demographic sorting's relationships with these variables are largely insignificant, stable or weakening over time, or both. The sole exception is vote turnout, but only when removing attitudinal sorting from the models.

Table 6: Partisan sorting's predictive power beyond affective polarization (summarized)

Type of sorting	Model	Activism (1972–2024)	Split-ticket voting (1972–2024)	Validated vote turnout (1980–1988, 2012–2016)	Party ID stability (1972, 1992, 2000, & 2016)	Party ID strength (1972–2024)
Demographic sorting	Multivariate	-0.04^{**} (0.01) Stable	-0.75^{**} (0.11) Stable	0.07 (0.13) Strengthening†	0.56^{*} (0.22) Weakening*	0.26^{**} (0.02) Stable
	Bivariate	0.04^{**} (0.01) Stable	-1.26^{**} (0.11) Stable	0.42^{**} (0.12) Strengthening†	1.37^{**} (0.21) Weakening†	0.37^{**} (0.02) Stable
Ideological sorting	Multivariate	0.14^{**} (0.01) Weakening*	-0.56^{**} (0.09) Stable	0.53^{**} (0.10) Stable	1.04^{**} (0.19) Stable	0.19^{**} (0.02) Strengthening**
	Bivariate	0.21^{**} (0.01) Stable	-1.23^{**} (0.08) Strengthening*	0.86^{**} (0.09) Stable	1.92^{**} (0.16) Stable	0.32^{**} (0.02) Strengthening**
Policy sorting	Multivariate	0.11^{**} (0.01) Stable	-1.08^{**} (0.11) Stable	0.63^{**} (0.13) Strengthening**	1.16^{**} (0.23) Strengthening†	0.19^{**} (0.02) Weakening†
	Bivariate	0.22^{**} (0.01) Stable	-1.79^{**} (0.10) Strengthening†	1.02^{**} (0.10) Strengthening**	2.29^{**} (0.17) Strengthening*	0.39^{**} (0.02) Stable
Group sentiment sorting	Multivariate	0.09^{**} (0.01) Stable	-0.69^{**} (0.12) Strengthening*	0.30^{*} (0.14) Stable	1.40^{**} (0.22) Stable	0.09^{**} (0.02) Stable
	Bivariate	0.21^{**} (0.01) Stable	-1.65^{**} (0.10) Strengthening**	0.90^{**} (0.11) Stable	2.56^{**} (0.19) Stable	0.35^{**} (0.02) Strengthening*
Economic attitude sorting	Multivariate	0.10^{**} (0.01) Stable	-0.78^{**} (0.12) Stable	0.53^{**} (0.13) Stable	0.89^{**} (0.23) Weakening**	0.21^{**} (0.02) Stable
	Bivariate	0.20^{**} (0.01) Stable	-1.61^{**} (0.11) Strengthening†	0.95^{**} (0.11) Strengthening*	2.02^{**} (0.19) Weakening†	0.40^{**} (0.02) Stable
Cultural attitude sorting	Multivariate	0.08^{**} (0.01) Stable	-0.67^{**} (0.14) Stable	0.64^{**} (0.15) Stable	1.66^{**} (0.27) Strengthening**	0.01 (0.03) Stable
	Bivariate	0.23^{**} (0.01) Stable	-1.69^{**} (0.12) Strengthening*	1.16^{**} (0.13) Stable	2.78^{**} (0.20) Strengthening**	0.32^{**} (0.02) Stable
Racial attitude sorting	Multivariate	0.05^{**} (0.01) Stable	-0.62^{**} (0.14) Stable	0.00 (0.14) Stable	0.80^{*} (0.28) Stable	0.09^{**} (0.03) Stable
	Bivariate	0.18^{**} (0.01) Stable	-1.59^{**} (0.13) Stable	0.63^{**} (0.12) Strengthening†	2.25^{**} (0.22) Stable	0.34^{**} (0.03) Stable

A summary of five Analysis #2 replications using behavior and partisan identity as dependent variables. Appendix O displays full results (pp. A.22–A.24). Activism and party ID models use multilinear regressions; the rest use probit regressions. Cells display α_s coefficients and standard errors (Equation #1), followed by the direction and significance of β_3 (Equation #3). $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

Interestingly, however, party ID strength's relationship with demographic sorting matches or exceeds its relationships with all forms of attitudinal sorting. This supports one of Mason's

(2018) important insights: When demographic affiliations align, they magnify partisanship's strength as a superordinate "mega-identity" (p. 14). This process still cannot explain over-time change—both demographic sorting and its relationship with party ID strength are stable over time—but it lends credence to prior scholarship on partisan social identity and suggests a promising avenue for future study.

Analyses #5a–5b explore the causal relationship between attitudinal sorting and affective polarization, but more work is needed here. In particular, panel studies with more than three waves—collected over shorter intervals—will provide more precise estimates of reciprocal causal effects. Experiments might assess whether an exogenous reduction in attitudinal sorting through persuasion reduces partisan animus.

This paper does not otherwise investigate causes of attitudinal sorting. Prior research suggests elite cues have been central in this process (Druckman et al. 2013; Lenz 2012), though other work attributes elite polarization to the electorate's earlier sorting on racial and cultural issues (Schickler 2016; O'Brian 2024). Future theories of partisan sorting and affective polarization should examine how mass attitudes shape elite incentives while elite rhetoric reconfigures mass opinion. Such work would benefit from attention to the nationalization of meso-institutions: Interest groups, state parties, and media organizations have increasingly homogenized within partisan coalitions (Pierson and Schickler 2024). This decline in cross-cutting institutional cleavages likely helps explain the loss of cross-cutting attitudinal cleavages.

CONCLUSION

Social groups remain critical to understanding our polarized politics, but demographic sorting itself is only a stable contributor to affective polarization and cannot explain its rise. The more important impact of groups comes from differences in identity-inflected attitudes, rather than basic group memberships. Distinguishing between the two is important, for many reasons.

First, distorted *perceptions* of the parties' demographic compositions lead to increased polarization (Ahler and Sood 2018). Therefore, just as media coverage of polarization can intensify partisan divisions (Levendusky and Malhotra 2016a), raising alarm about demographic polarization could unjustifiably fuel partisan animus. Second, attitudinal sorting may better explain racial depolarization. Republican gains among non-white voters reflect a combination of education polarization and ideological sorting (Fraga et al. 2025; Ruffini 2023), suggesting even demographic shifts in party coalitions can trace back to attitudes. Furthermore, stable demographic sorting amid rising racial attitude sorting highlights a key distinction between polarization "by race" and "about race" (Sides and Tesler 2024).

Finally, demographic and attitudinal sorting present different challenges for American democracy. Those concerned about demographic sorting often warn that voters are driven by tribalism, even though they "disrespectfully agree" on policy matters (Mason 2015). By contrast, polarization driven by attitudinal sorting reflects real differences in what each side wants from government and society. This may be a cause for pessimism: Such polarization is unlikely to fade as social groups realign within the parties—if those realignments are based on substantive disagreements.

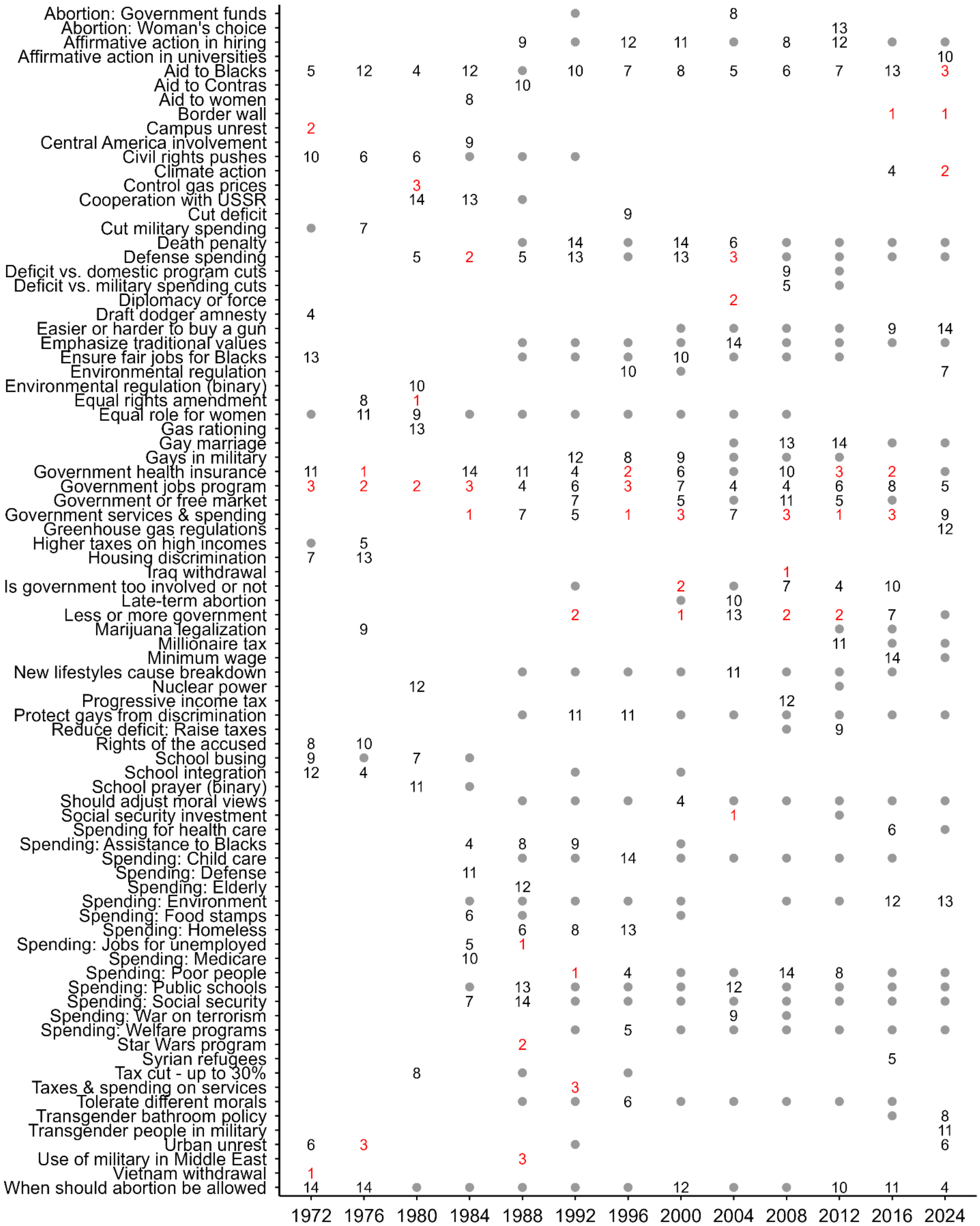
However, if Americans are more divided by what they think than by who they are, this could also mean polarization is less inevitable than many assume. Changing demographics is harder than changing attitudes via persuasion (Coppock 2023). Thus, attitudinal sorting could be more reversible than demographic sorting, which in turn suggests attitude-driven polarization is more reversible—or at least that Americans can choose whether to engage in the deliberation and consensus-building necessary to reverse it. To this end, compromise and accommodation—even on identity-inflected issues—are far from impossible (see Rauch 2015; Yglesias 2023). Divisions based on the mere existence of different groups offer no comparable means of resolution.

Thus, observers may overestimate how intractable America's identity-inflected divide actually is. The country is less divided by inevitable group-based animus spurred by increased diversity and more by differences of opinion on an overarching question: "How much should America focus on restorative justice for marginalized groups, relative to other priorities?" Though surely difficult, finding an acceptable consensus may seem less impossible if we stop assuming demographic traits are sufficient stand-ins for more malleable attitudes.

APPENDIX: ITEMS CHOSEN FOR SORTING MEASURES

Attitudinal sorting measures use the eight or 14 items that best predict partisanship each year, as determined by random forest importance scores. Figures A1–A5 list all items selected at least once.

Figure A1: Items used to measure policy sorting



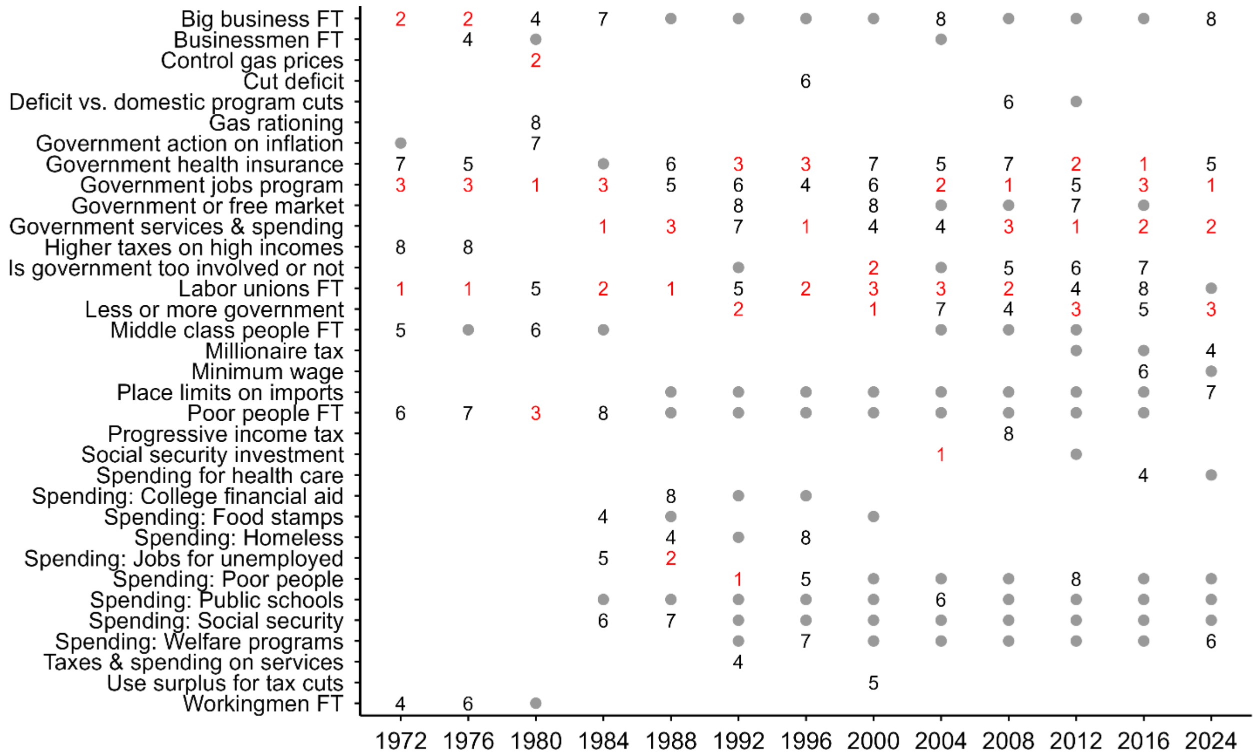
Sorting was measured using each year's strongest predictors of partisanship (1=strongest chosen predictor; 14=weakest chosen predictor). Circles mark available items that were not chosen.

Figure A2: Items used to measure group sentiment sorting

	1972	1976	1980	1984	1988	1992	1996	2000	2004	2008	2012	2016	2024
Anti-abortionists FT				14	●								
Asian Americans FT						●		●	14	●	●	●	●
Big business FT	1	1	2	6	10	5	11	8	5	●	●	8	11
Black Lives Matter FT												1	1
Black militants FT	●	8	7	5									
Blacks FT	7	7	3	●	3	7	10	6	7	3	2	10	13
Businessmen FT		3	●						9				
Catholics FT	●	13		●	●	14		5	●	●	7		
Christian fundamentalists FT					●	12	8	13	8	13	12	7	5
Christians FT										●	14	12	7
Civil rights leaders FT	3	4	1	7	2								
Elderly/older people FT		●	●	11	8		12	●	●				
Environmentalists FT			●		11	8	3	2	4	2			
Evangelicals FT			8	8	●								
Farmers FT	●		13										
Feminists FT					4	1		4	2	8	6	3	4
Gays and lesbians FT				●	12	4	4	11	6	12	●	11	9
Hispanics FT		9	●	●	13	●	14	●	12	●	10	13	●
Illegal immigrants FT					7	●			10	5	4	4	2
Israel FT									●	6			
Jews FT	●	●			●	●		●	●	●		14	14
Labor unions FT	2	2	5	1	1	2	2	1	1	1	1	2	10
Lawyers FT						11							
Middle class people FT	●	●	14	9					●	14	●		
Ministers who lead protests FT	4												
Mormons FT											5		
Muslims FT									●	11	8	5	12
People on welfare FT		6	4	10	9	10	9	9	●	7	3		
Police FT	8	11				13						9	6
Poor people FT	14	10	6	3	6	9	7	●	●	10	13	●	
Protestants FT	11	14						10					
Radical students FT	5	5	10										
Rural Americans FT													8
Southerners FT	●	●	11			●			13	●			
The Christian Coalition FT							5	12					
The military FT	13	●	12	4	5	6	6	7	3	4	9		
Transgender people FT												6	3
Urban rioters FT	6												
Wealthy/rich people FT									11	●	11	●	
Whites FT	●	●	●	13	14	●	13	14	●	●	●	●	●
Women's movement FT	12	●	9	2		3	1	3					
Women FT		●		12	●				●				
Working class people FT									●	9	●		
Workingmen FT	9	12	●										
Young people FT	10	●	●						●				

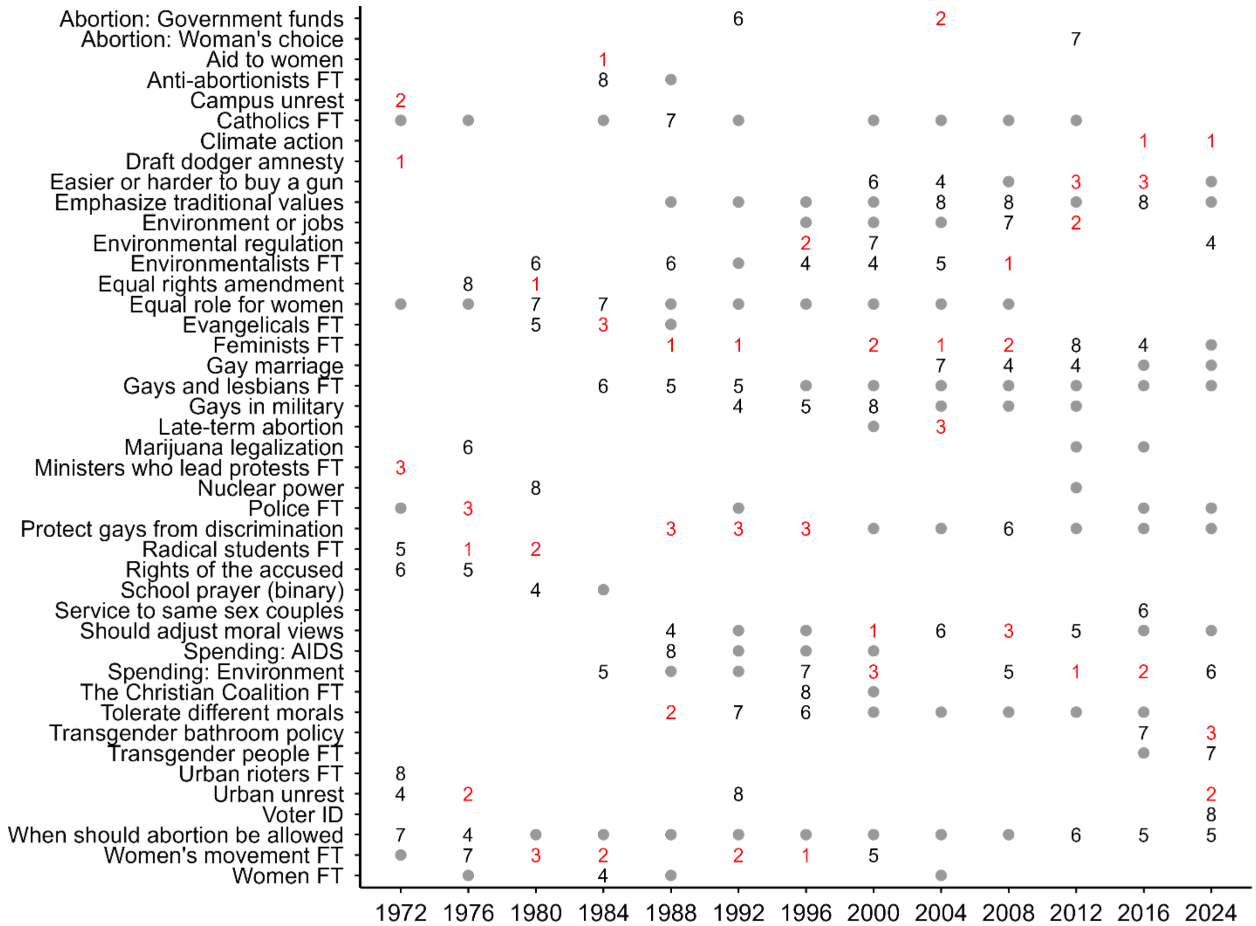
Sorting was measured using each year's strongest predictors of partisanship (1=strongest chosen predictor; 14=weakest chosen predictor). Circles mark available items that were not chosen. FT = feeling thermometer.

Figure A3: Items used to measure economic attitude sorting



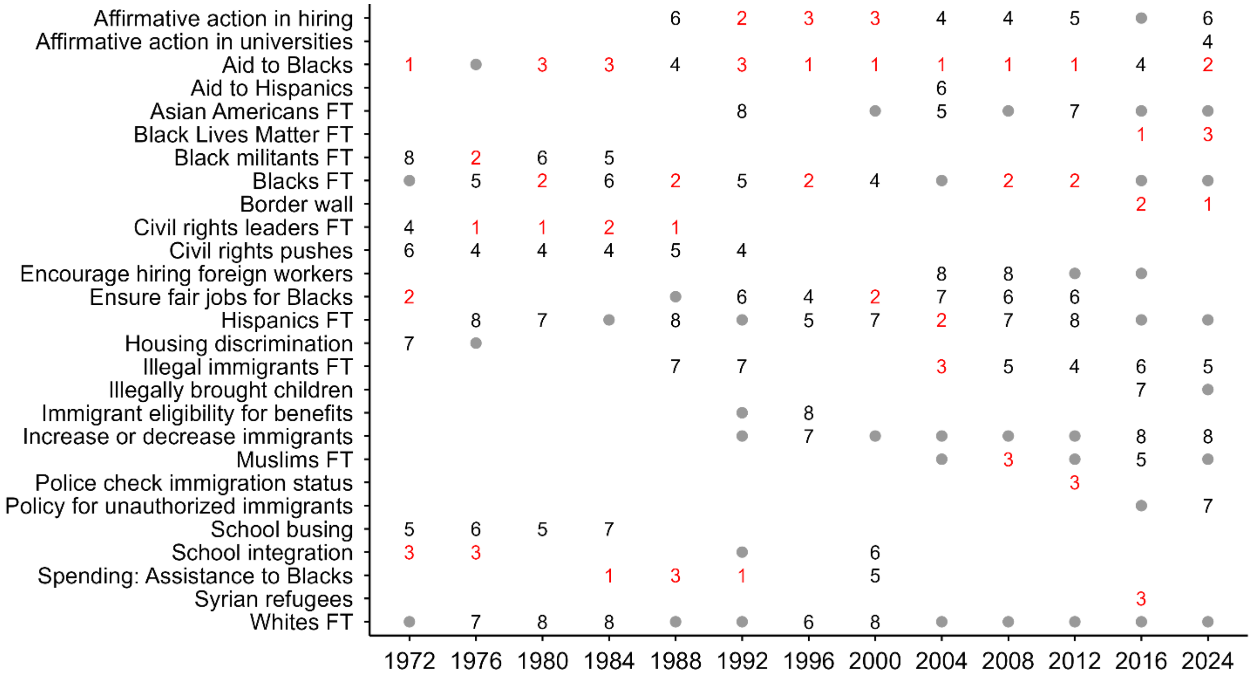
Sorting was measured using each year's strongest predictors of partisanship (1=strongest chosen predictor; 8=weakest chosen predictor). Circles mark available items that were not chosen. FT = feeling thermometer.

Figure A4: Items used to measure cultural attitude sorting



Sorting was measured using each year's strongest predictors of partisanship (1=strongest chosen predictor; 8=weakest chosen predictor). Circles mark available items that were not chosen. FT = feeling thermometer.

Figure A5: Items used to measure racial attitude sorting



Sorting was measured using each year's strongest predictors of partisanship (1=strongest chosen predictor; 8=weakest chosen predictor). Circles mark available items that were not chosen. FT = feeling thermometer.

REFERENCES

Abramowitz, Alan I. 1994. "Issue Evolution Reconsidered: Racial Attitudes and Partisanship in the U.S. Electorate." *American Journal of Political Science* 38 (1): 1–24.

Abramowitz, Alan I. 2011. *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy.* New Haven, CT: Yale University Press.

Abramowitz, Alan I., and Kyle L. Saunders. 2008. "Is Polarization a Myth?" *Journal of Politics* 70 (2): 542–555.

Achen, Christopher H., and Larry M. Bartels. 2017. *Democracy for Realists: Why Elections Do Not Produce Responsive Government.* Princeton: Princeton University Press.

Agadjanian, Alexander, and Dean Lacy. 2021. "Changing Votes, Changing Identities? Racial Fluidity and Vote Switching in the 2012–2016 U.S. Presidential Elections." *Public Opinion Quarterly* 85 (3): 737–752.

Agar, Herbert. 1950. *The Price of Union.* Boston, MA: Houghton Mifflin.

Ahler, Douglas J., and Gaurav Sood. 2018. "The Parties in Our Heads: Misperceptions about Party Composition and Their Consequences." *Journal of Politics* 80 (3): 964–981.

American National Election Studies. 2026. *ANES Time Series Cumulative Data File* [Dataset and documentation]. February 5, 2026 version. https://electionstudies.org/data-center/anes-time-series-cumulative-data-file/.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder. 2006. "Purple America." *Journal of Economic Perspectives* 20 (2): 97–118.

Bafumi, Joseph, and Robert Y. Shapiro. 2009. "A New Partisan Voter." *Journal of Politics* 71 (1): 1–24.

Bailey, Michael A. 2024. *Polling at a Crossroads: Rethinking Modern Survey Research.* Cambridge: Cambridge University Press.

Baldassarri, Delia, and Andrew Gelman. 2008. "Partisans without Constraint: Political Polarization and Trends in American Public Opinion." *American Journal of Sociology* 114 (2): 408–446.

Baldassarri, Delia, and Jona de Jong. 2025. "Partisanship Meets Social Networks: How Politically Heterogeneous Acquaintances and Close Relationships Buffer Partisan Animosity." *Sociological Science* 12: 409–430.

Barreto, Matt, and Gary M. Segura. 2014. *Latino America: How America's Most Dynamic Population Is Poised to Transform the Politics of the Nation.* New York: PublicAffairs.

Bartels, Larry M. 2006. "What's the Matter with 'What's the Matter with Kansas?'" *Quarterly Journal of Political Science* 1: 201–226.

Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.

Bougher, Lori D. 2017. "The Correlates of Discord: Identity, Issue Alignment, and Political Hostility in Polarized America." *Political Behavior* 39: 731–762.

Brader, Ted, Joshua A. Tucker, and Andrew Therriault. 2014. "Cross Pressure Scores: An Individual-Level Measure of Cumulative Partisan Pressures Arising from Social Group Memberships." *Political Behavior* 36: 23–51.

Broockman, David E. 2016. "Approaches to Studying Policy Representation." *Legislative Studies Quarterly* 41 (1): 181–215.

Broockman, David E., Joshua L. Kalla, and Sean J. Westwood. 2023. "Does Affective Polarization Undermine Democratic Norms or Accountability? Maybe Not." *American Journal of Political Science* 67 (3): 808–828.

Brubaker, Rogers. 2002. "Ethnicity without Groups." *European Journal of Sociology* 43 (2): 163–189.

Calvo, Richard, Vincent Pons, and Jesse M. Shapiro. 2024. "Pitfalls of Demographic Forecasts of U.S. Elections." NBER Working Paper No. 33016. http://www.nber.org/papers/w33016.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. [1960] 1980. *The American Voter.* Chicago: University of Chicago Press.

Campbell, James E. 2016. *Polarized: Making Sense of a Divided America.* Princeton: Princeton University Press.

Campos, Nicolas, and Christopher Federico. 2026. "A New Measure of Affective Polarization." *American Political Science Review* 120 (1): 160–178.

Carsey, Thomas M., and Geoffrey C. Layman. 2006. "Changing Sides or Changing Minds? Party Identification and Policy Preferences in the American Electorate." *American Journal of Political Science* 50 (2): 464–477.

Cavari, Amnon, and Guy Freedman. 2018. "Polarized Mass or Polarized Few? Assessing the Parallel Rise of Survey Nonresponse and Measures of Polarization." *Journal of Politics* 80 (2): 719–725.

Cavari, Amnon, and Guy Freedman. 2023. "Survey Nonresponse and Mass Polarization: The Consequences of Declining Contact and Cooperation Rates." *American Political Science Review* 117 (1): 332–339.

Chotiner, Isaac. 2024. "What the Closeness of This Election Suggests About the Future of American Politics." *The New Yorker*, October 17. https://www.newyorker.com/news/q-and-a/what-the-closeness-of-this-election-suggests-about-the-future-of-american-politics.

Clifford, Scott, Elizabeth Simas, and JeongKyu Suh. 2025. "The Policy Basis of Group Sentiments." *Political Science Research and Methods* 13 (2): 482–488.

Conover, Pamela J. 1988. "The Role of Social Groups in Political Thinking." *British Journal of Political Science* 18: 51–76.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter, 206–261. New York: Free Press.

Coppock, Alexander. 2023. *Persuasion in Parallel: How Information Changes Minds about Politics.* Chicago: University of Chicago Press.

Dahl, Robert. 1981. *Democracy in the United States: Promise and Performance.* Boston, MA: Houghton Mifflin.

Dias, Nicholas C., and Yphtach Lelkes. 2022. "The Nature of Affective Polarization: Disentangling Policy Disagreement from Partisan Identity." *American Journal of Political Science* 66 (3): 775–790.

Dias, Nicholas C., Yphtach Lelkes, and Jacob Pearl. 2025. "American Partisans Vastly Under-Estimate the Diversity of Other Partisans' Policy Attitudes." *Political Science Research and Methods* 13 (3): 725–735.

Dowling, Conor. 2019. "CCES 2016, Team Module of University of Mississippi (OLE MISS)." *Harvard Dataverse.* https://doi.org/10.7910/DVN/0JIMKY.

Druckman, James N., Donald P. Green, and Shanto Iyengar. 2023. "Does Affective Polarization Contribute to Democratic Backsliding in America?" *ANNALS of the American Academy of Political and Social Science* 708 (1): 137–163.

Druckman, James N., Erik Peterson, and Rune Slothuus. 2013. "How Elite Partisan Polarization Affects Public Opinion Formation." *American Political Science Review* 107 (1): 57–79.

Druckman, James N., and Matthew S. Levendusky. 2019. "What Do We Measure When We Measure Affective Polarization?" *Public Opinion Quarterly* 83 (1): 114–122.

Druckman, James N., Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. 2022. "(Mis)estimating Affective Polarization." *Journal of Politics* 84 (2): 1106–1117.

Druckman, James N., Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. 2024. *Partisan Hostility and American Democracy: Explaining Political Divisions and When They Matter.* Chicago: University of Chicago Press.

Dyck, Joshua J., and Shanna Pearson-Merkowitz. 2023. *The Power of Partisanship.* Oxford: Oxford University Press.

Egan, Patrick J. 2020. "Identity as Dependent Variable: How Americans Shift Their Identities to Align with Their Politics." *American Journal of Political Science* 64 (3): 699–716.

Elder, Elizabeth Mitchell, and Neil A. O'Brian. 2022. "Social Groups as the Source of Political Belief Systems: Fresh Evidence on an Old Theory." *American Political Science Review* 116 (4): 1407–1424.

Ellis, Christopher, and James A. Stimson. 2012. *Ideology in America.* Cambridge: Cambridge University Press.

Finkel, Eli J., Christopher A. Bail, Mina Cikara, Peter H. Ditto, Shanto Iyengar, Samara Klar, Lilliana Mason, et al. 2020. "Political Sectarianism in America." *Science* 370 (6516): 533–536.

Finkel, Steven E. 1995. *Causal Analysis with Panel Data.* Thousand Oaks, CA: Sage Publications.

Fiorina, Morris P., Samuel J. Abrams, and Jeremy C. Pope. 2005. *Culture War? The Myth of a Polarized America.* New York: Pearson Longman.

Fowler, Anthony. 2020. “Partisan Intoxication or Policy Voting?” *Quarterly Journal of Political Science* 15 (2): 141–179.

Fowler, Anthony, Gregory A. Huber, Rongbo Jin, and Lilla V. Orr. 2026. “How Do Americans Explain Their Party Identification and Out-Partisan Animosity?” *Public Opinion Quarterly*, nfag044. https://doi.org/10.1093/poq/nfag044.

Fraga, Bernard L., Yamil R. Velez, and Emily A. West. 2025. “Reversion to the Mean, or Their Version of the Dream? Latino Voting in an Age of Populism.” *American Political Science Review* 119 (1): 517–525.

French, David. 2020. *Divided We Fall: America’s Secession Threat and How to Restore Our Nation.* New York: St. Martin’s Press.

Gidron, Noam, James Adams, Will Horn, and Thomas Tichelbaecker. 2025. “Beyond Observational Relationships: Evidence from a Ten-Country Experiment that Policy Disputes Cause Affective Polarization.” *British Journal of Political Science* 55 (e84): 1–10.

Gilens, Martin. 1999. *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy.* Chicago: University of Chicago Press.

Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters.* New Haven, CT: Yale University Press.

Grossmann, Matt, and David A. Hopkins. 2024. *Polarized by Degrees: How the Diploma Divide and the Culture War Transformed American Politics.* Cambridge: Cambridge University Press.

Guntermann, Eric. 2022. “Socially-Sorted Partisans Are More Affectively Polarized: Evidence from the United States and Elsewhere.” Working Paper.

Hamaker, Ellen L., Rebecca M. Kuiper, and Raoul P. P. P. Grasman. 2015. “A Critique of the Cross-Lagged Panel Model.” *Psychological Methods* 20 (1): 102–116.

Hare, Christopher, and Mikayla Kutsuris. 2023. “Measuring Swing Voters with a Supervised Machine Learning Ensemble.” *Political Analysis* 31 (4): 537–553.

Harteveld, Eelco. 2021. “Ticking All the Boxes? A Comparative Study of Social Sorting and Affective Polarization.” *Electoral Studies* 72: 102337.

Highton, Benjamin, and Cindy D. Kam. 2011. “The Long-Term Dynamics of Partisanship and Issue Orientations.” *Journal of Politics* 73 (1): 202–215.

Hillygus, D. Sunshine, and Todd G. Shields. 2008. *The Persuadable Voter: Wedge Issues in Presidential Campaigns.* Princeton: Princeton University Press.

Horan, Patrick M. 1971. “Social Positions and Political Cross-Pressures: A Re-Examination.” *American Sociological Review* 36 (4): 650–660.

Huddy, Leonie. 2001. “From Social to Political Identity: A Critical Examination of Social Identity Theory.” *Political Psychology* 22 (1): 127–156.

Huddy, Leonie, and Omer Yair. 2021. “Reducing Affective Polarization: Warm Group Relations or Policy Compromise?” *Political Psychology* 42 (2): 291–309.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. “Affect, Not Ideology: A Social Identity Perspective on Polarization.” *Public Opinion Quarterly* 76 (3): 405–431.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2019. “The Origins and Consequences of Affective Polarization in the United States.” *Annual Review of Political Science* 22: 129–146.

Jacobs, Nicholas F., and Daniel M. Shea. 2023. *The Rural Voter: The Politics of Place and the Disuniting of America.* New York: Columbia University Press.

Jardina, Ashley. 2019. *White Identity Politics.* Cambridge: Cambridge University Press.

Jennings, M. Kent, Gregory B. Markus, Richard G. Niemi, and Laura Stoker. 2005. *Youth-Parent Socialization Panel Study, 1965–1997: Four Waves Combined.* Inter-university Consortium for Political and Social Research [distributor], November 4. https://doi.org/10.3886/ICPSR04037.v1.

Jin, Rongbo, Samara Klar, Fabian G. Neuner, and Mark Ramirez. 2025. "How Cross-Cutting Ties Reduce Affective Polarization: Evidence from Latino Americans." *Political Behavior* 47: 1925–1947.

Jones, Robert P. 2016. *The End of White Christian America.* New York: Simon & Schuster.

Kalmoe, Nathan P., and Lilliana Mason. 2022. *Radical American Partisanship: Mapping Violent Hostility, Its Causes, and the Consequences for Democracy.* Chicago: University of Chicago Press.

Kane, John V., Lilliana Mason, and Julie Wronski. 2021. "Who's at the Party? Group Sentiments, Knowledge, and Partisan Identity." *Journal of Politics* 83 (4): 1783–1799.

Keith, Bruce E., David B. Magleby, Candice J. Nelson, Elizabeth Orr, Mark C. Westlye, and Raymond E. Wolfinger. 1992. *The Myth of the Independent Voter.* Berkeley, CA: University of California Press.

Kim, Seo-young Silvia, and Jan Zilinsky. 2024. "Division Does Not Imply Predictability: Demographics Continue to Reveal Little about Voting and Partisanship." *Political Behavior* 46: 67–87.

Kinder, Donald R., and Lynn M. Sanders. 1996. *Divided by Color: Racial Politics and Democratic Ideals.* Chicago: University of Chicago Press.

Klein, Ezra. 2020. *Why We're Polarized.* New York: Avid Reader Press.

Kollman, Ken, and John E. Jackson. 2026. "Revisiting Ideology Measures in Political Surveys." *Journal of Politics* 88 (1): 1–17.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1944. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign.* New York: Columbia University Press.

Lelkes, Yphtach. 2018. "Affective Polarization and Ideological Sorting: A Reciprocal, Albeit Weak, Relationship." *Forum: A Journal of Applied Research in Contemporary Politics* 16 (1): 67–79.

Lelkes, Yphtach. 2021. "Policy over Party: Comparing the Effects of Candidate Ideology and Party on Affective Polarization." *Political Science Research and Methods* 9 (1): 189–196.

Lenz, Gabriel S. 2012. *Follow the Leader? How Voters Respond to Politicians' Policies and Performance.* Chicago: University of Chicago Press.

Levendusky, Matthew. 2009. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans.* Chicago: University of Chicago Press.

Levendusky, Matthew. 2023. *Our Common Bonds: Using What Americans Share to Help Bridge the Partisan Divide.* Chicago: University of Chicago Press.

Levendusky, Matthew, and Neil Malhotra. 2016a. "Does Media Coverage of Partisan Polarization Affect Political Attitudes?" *Political Communication* 33 (2): 283–301.

Levendusky, Matthew, and Neil Malhotra. 2016b. "(Mis)perceptions of Partisan Polarization in the American Public." *Public Opinion Quarterly* 80 (S1): 378–391.

Lipset, Seymour Martin. 1960. *Political Man: The Social Bases of Politics.* Garden City, NY: Doubleday.

Lucas, Richard E. 2023. “Why the Cross-Lagged Panel Model Is Almost Never the Right Choice.” *Advances in Methods and Practices in Psychological Science* 6 (1): 1–22.

Marble, William. 2024. “What Explains Educational Realignment Among White Americans?” Working Paper. https://williammarble.co/docs/EducPolarization.pdf.

Margolis, Michele F. 2018. “How Politics Affects Religion: Partisanship, Socialization, and Religiosity in America.” *Journal of Politics* 80 (1): 30–43.

Mason, Lilliana. 2015. “‘I Disrespectfully Agree’: The Differential Effects of Partisan Sorting on Social and Issue Polarization.” *American Journal of Political Science* 59 (1): 128–145.

Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity*. Chicago: University of Chicago Press.

Mason, Lilliana, and Julie Wronski. 2018. “One Tribe to Bind Them All: How Our Social Group Attachments Strengthen Partisanship.” *Political Psychology* 39 (S1): 257–277.

Mason, Lilliana, Julie Wronski, and John V. Kane. 2021. “Activating Animus: The Uniquely Social Roots of Trump Support.” *American Political Science Review* 115 (4): 1508–1516.

Mellon, Jonathan, and Christopher Prosser. 2021. “Correlation with Time Explains the Relationship between Survey Nonresponse and Mass Polarization.” *Journal of Politics* 83 (1): 390–395.

Mellon, Jonathan, and Christopher Prosser. 2025. “Regularized Regression Can Reintroduce Backdoor Confounding: The Case of Mass Polarization.” *American Political Science Review* 119 (4): 2002–2010.

Miller, Arthur H., and Christopher Wlezien. 1993. “The Social Group Dynamics of Partisan Evaluations.” *Electoral Studies* 12 (1): 5–22.

Montgomery, Jacob M., and Santiago Olivella. 2018. “Tree-Based Models for Political Science Data.” *American Journal of Political Science* 62 (3): 729–744.

Mutz, Diana C. 2006. *Hearing the Other Side: Deliberative versus Participatory Democracy.* Cambridge: Cambridge University Press.

Newman, Benjamin, Tyler Reny, and Marcel Roman. 2026. “The Immigration-to-Reproduction Shift: Latino Population Growth and White Support for Legal Abortion.” Working Paper. https://www.marcelroman.com/pdfs/wps/abortion.pdf.

O’Brian, Neil. 2024. *The Roots of Polarization: From the Racial Realignment to the Culture Wars.* Chicago: University of Chicago Press.

Orr, Lilla V., Anthony Fowler, and Gregory A. Huber. 2023. “Is Affective Polarization Driven by Identity, Loyalty, or Substance?” *American Journal of Political Science* 67 (4): 948–962.

Orr, Lilla V., and Gregory A. Huber. 2020. “The Policy Basis of Measured Partisan Animosity in the United States.” *American Journal of Political Science* 64 (3): 569–586.

Pierson, Paul, and Eric Schickler. 2024. *Partisan Nation: The Dangerous New Logic of American Politics in a Nationalized Era.* Chicago: University of Chicago Press.

Ramirez, Mark D., and David A. M. Peterson. 2020. *Ignored Racism: White Animus Toward Latinos.* Cambridge: Cambridge University Press.

Rauch, Jonathan. 2015. *Political Realism: How Hacks, Machines, Big Money, and Back-Room Deals Can Strengthen American Democracy.* Washington, DC: Brookings Institution Press.

Rawlings, Craig M. 2022. “Becoming an Ideologue: Social Sorting and the Microfoundations of Polarization.” *Sociological Science* 9 (13): 313–345.

Robison, Joshua, and Rachel L. Moskowitz. 2019. “The Group Basis of Partisan Affective Polarization.” *Journal of Politics* 81 (3): 1075–1079.

Rodden, Jonathan A. 2019. *Why Cities Lose: The Deep Roots of the Urban-Rural Political Divide.* New York: Basic Books.

Rogowski, Jon C., and Joseph L. Sutherland. 2016. “How Ideology Fuels Affective Polarization.” *Political Behavior* 38: 485–508.

Rothschild, Jacob E., Adam J. Howat, Richard M. Shafranek, and Ethan C. Busby. 2019. “Pigeonholing Partisans: Stereotypes of Party Supporters and Partisan Polarization.” *Political Behavior* 41: 423–443.

Ruffini, Patrick. 2023. *Part of the People: Inside the Multiracial Populist Coalition Remaking the GOP.* New York: Simon & Schuster.

Sasse, Ben. 2018. *Them: Why We Hate Each Other—and How to Heal.* New York: St. Martin’s Press.

Schickler, Eric. 2016. *Racial Realignment: The Transformation of American Liberalism, 1932–1965.* Princeton: Princeton University Press.

Sears, David O. 1993. “Symbolic Politics: A Socio-Psychological Theory.” In *Explorations in Political Psychology*, eds. Shanto Iyengar and William J. McGuire, 113–149. Durham, NC: Duke University Press.

Sides, John, Chris Tausanovitch, and Lynn Vavreck. 2022. *The Bitter End: The 2020 Presidential Campaign and the Challenge to American Democracy.* Princeton: Princeton University Press.

Sides, John, and Michael Tesler. 2024. “America Is Less Polarized by Race—but More Polarized about Race.” *Good Authority*, April 3. https://goodauthority.org/news/america-is-less-polarized-by-race-but-more-polarized-about-race/.

Sides, John, Michael Tesler, and Lynn Vavreck. 2018. *Identity Crisis: The 2016 Presidential Campaign and the Battle for the Meaning of America.* Princeton: Princeton University Press.

Simas, Elizabeth N. 2023. *In Defense of Ideology: Reexamining the Role of Ideology in the American Electorate.* Cambridge: Cambridge University Press.

Tajfel, Henri, and John C. Turner. 1986. “The Social Identity Theory of Intergroup Behavior.” In *Psychology of Intergroup Relations*, eds. Stephen Worchel and William G. Austin, 7–24. Chicago: Nelson-Hall.

Tausanovitch, Chris, Lynn Vavreck, Tyler Reny, Alex Rossell Hayes, and Aaron Rudkin. 2019. “Democracy Fund + UCLA Nationscape Methodology and Representativeness Assessment.” https://www.voterstudygroup.org/uploads/reports/Data/NS-Methodology-Representativeness-Assessment.pdf.

Tesler, Michael. 2016. *Post-Racial or Most-Racial? Race and Politics in the Obama Era.* Chicago: University of Chicago Press.

Thornton, Judd R. 2025. “The Impact of Survey Mode on the Measurement of Polarization in the Public: 2016–2020 U.S. Panel Evidence.” *American Politics Research* 53 (5): 531–543.

Tyler, Matthew, and Shanto Iyengar. 2023. “Learning to Dislike Your Opponents: Political Socialization in the Era of Polarization.” *American Political Science Review* 117 (1): 347–354.

Tyler, Matthew, and Shanto Iyengar. 2024. "Testing the Robustness of the ANES Feeling Thermometer Indicators of Affective Polarization." *American Political Science Review* 118 (3): 1570–1576.

van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2011. "mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45 (3): 1–67.

Voelkel, Jan G., James Chu, Michael N. Stagnaro, Joseph S. Mernyk, Chrystal Redekopp, Sophia L. Pink, James N. Druckman, David G. Rand, and Robb Willer. 2023. "Interventions Reducing Affective Polarization Do Not Necessarily Improve Anti-Democratic Attitudes." *Nature Human Behaviour* 7 (1): 55–64.

Voelkel, Jan G., Michael N. Stagnaro, James Y. Chu, Sophia L. Pink, Joseph S. Mernyk, Chrystal Redekopp, Isaias Ghezae, et al. 2024. "Megastudy Testing 25 Treatments to Reduce Antidemocratic Attitudes and Partisan Animosity." *Science* 386 (6719): eadh4764.

Wattenberg, Martin P. 2019. "The Changing Nature of Mass Belief Systems: The Rise of Concept and Policy Ideologues." *Critical Review* 31 (2): 198–229.

Weber, Christopher, and Samara Klar. 2019. "Exploring the Psychological Foundations of Ideological and Social Sorting." *Political Psychology* 40 (S1): 215–243.

Webster, Steven W., and Alan I. Abramowitz. 2017. "The Ideological Foundations of Affective Polarization in the U.S. Electorate." *American Politics Research* 45 (4): 621–647.

Westwood, Sean J., and Erik Peterson. 2022. "The Inseparability of Race and Partisanship in the United States." *Political Behavior* 44: 1125–1147.

White, Ismail K., and Chryl N. Laird. 2020. *Steadfast Democrats: How Social Forces Shape Black Political Behavior.* Princeton: Princeton University Press.

Yeung, Eddy S. F., and Kai Quek. 2025. "Self-Reported Political Ideology." *Political Science Research and Methods* 13 (2): 412–433.

Yglesias, Matthew. 2023. "The Two Kinds of Progressives." *Slow Boring*, August 21. https://www.slowboring.com/p/the-two-kinds-of-progressives.

Zhirkov, Kirill, and Nicholas A. Valentino. 2022. "The Origins and Consequences of Racialized Schemas about U.S. Parties." *Journal of Race, Ethnicity, and Politics* 7 (3): 484–504.

Zingher, Joshua N. 2022. "Diploma Divide: Educational Attainment and the Realignment of the American Electorate." *Political Research Quarterly* 75 (2): 263–277.

Online Appendix

The Rise of Affective Polarization: Is It Who We Are, or What We Think?

American Journal of Political Science

Contents

Appendix A: Validating random forest sorting measurement A.0
Appendix B: Alternate measures of demographic sorting A.2
Appendix C: Excluding independents from sorting models A.4
Appendix D: Assessing MICE imputation's effect on sorting A.4
Appendix E: Online ANES time series trends A.5
Appendix F: Nonresponse bias and rising attitudinal sorting A.7
Appendix G: The influence of early demographic sorting A.10
Appendix H: Binning by year in Analysis #2 regressions A.13
Appendix I: Decompositions across several specifications A.13
Appendix J: Results when using same attitudes over time A.14
Appendix K: Analysis #4 replication details A.17
Appendix L: Items used to calculate sorting in panels A.18
Appendix M: A discussion of CLPMs and RI-CLPMs A.19
Appendix N: Studies used in experimental meta-analysis A.21
Appendix O: Beyond affective polarization A.22
References A.25

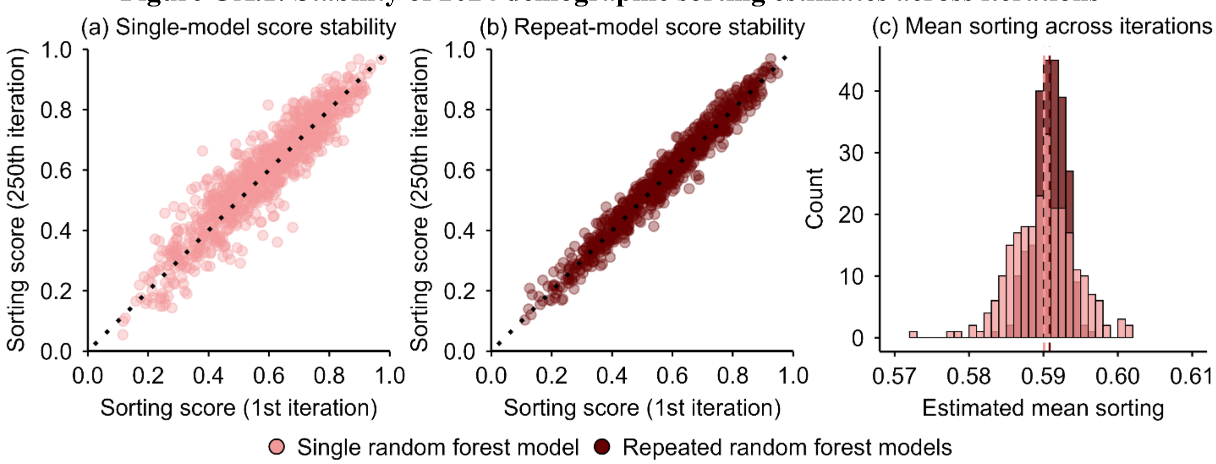
Appendix A: Validating random forest sorting measurement

Random forest models excel at prediction because of their ability to automatically account for nonlinear relationships and interactions between variables. As a result, they have become increasingly common in social science research (see Muchlinski et al. 2016; Bonica 2018; Peyton 2020; Kaufman and Rogowski 2024). This appendix addresses readers who remain skeptical of these models due to their eponymous randomness and black-box nature.

First, the sorting scores are quite stable across iterations—especially when using repeated models. I calculate 500 sets of demographic sorting scores for 2024: 250 iterations use repeated cross-validated models (as in the main text), and 250 use a single cross-validated model (for comparison). Single models provide impressive stability (correlating at 0.905 across the 250 scores), but repeated models achieve a near-perfect average correlation of 0.980. Furthermore, the range of scores each respondent receives across all iterations is, on average, only 0.140 when using repeated models (compared to 0.320 for single models).

Panels (a) and (b) in Figure OA.1 visualize this improvement: Single models score some respondents quite differently across iterations due to random chance, but repeated models prevent this. Panel (c) demonstrates how this improves aggregate estimates: All mean sorting level estimates are within 0.007 when using repeated models (whereas 9.2% of single-model estimates are outside this range). These results indicate that repeated random forest models measure partisan sorting stably.

Figure OA.1: Stability of 2024 demographic sorting estimates across iterations



Panels (a) and (b) plot the first and final sets of demographic sorting scores when using a single random forest model versus five pooled models, respectively. Panel (c) shows the distribution of 2024 "mean demographic sorting" estimates across both sets of 250 iterations. The dashed lines mark each set's mean estimate (0.590 for single models; 0.591 for repeated models).

Next, I show that this approach measures sorting reliably. Table OA.1 charts the demographic make-up of the most and least sorted respondents of each party at the start, middle, and end of the ANES time series. The results align with expectations, given knowledge of the parties' demographic compositions over time. For instance, well-sorted Democrats in 1952–1956 are likelier than poorly sorted Democrats to be southern, Catholic, and non-college-educated. In 2020–2024, well-sorted Democrats are more likely to be non-white, secular, college-educated, and urban. This suggests random forests capture salient cleavages and accurately measure partisan sorting.

Table OA.1: Demographics of least- and most-sorted partisans, over time

Demographic	Democrats						Republicans					
	1952–1956		1986–1990		2020–2024		1952–1956		1986–1990		2020–2024	
	Least	Most	Least	Most	Least	Most	Least	Most	Least	Most	Least	Most
White	100%	85%	99%	19%	98%	30%	83%	100%	35%	99%	41%	98%
Female	64%	46%	51%	65%	46%	64%	47%	57%	60%	43%	57%	44%
Protestant	99%	61%	81%	72%	59%	31%	61%	99%	55%	87%	26%	62%
Catholic	1%	28%	13%	18%	23%	11%	30%	1%	34%	10%	14%	18%
Attends church weekly	41%	48%	29%	32%	19%	10%	46%	44%	25%	38%	6%	34%
Never attends church	3%	5%	21%	15%	29%	63%	3%	3%	18%	11%	65%	19%
35 or younger	29%	32%	34%	28%	17%	37%	38%	16%	30%	32%	43%	14%
65 or older	13%	8%	15%	21%	32%	15%	7%	21%	22%	12%	10%	24%
Low income	30%	37%	11%	52%	22%	28%	43%	23%	53%	7%	18%	18%
High income	49%	31%	56%	16%	41%	44%	26%	59%	18%	62%	46%	48%
High school or less	56%	80%	27%	70%	41%	26%	78%	45%	75%	24%	29%	41%
Bachelor's degree +	14%	3%	35%	12%	29%	50%	7%	22%	11%	41%	45%	24%
Married	76%	81%	80%	44%	75%	36%	82%	82%	48%	86%	38%	81%
Northeastern	30%	16%	13%	19%	10%	20%	15%	47%	12%	12%	18%	10%
Southern	0%	54%	18%	57%	45%	37%	57%	0%	60%	13%	33%	51%
Rural	51%	38%	40%	28%	66%	16%	39%	41%	29%	33%	20%	68%
Urban	16%	36%	8%	48%	7%	56%	36%	16%	42%	7%	42%	8%
N	120	606	209	875	411	1668	148	348	177	475	312	1552

This table lists the percentage of Democrats and Republicans with demographic sorting scores above 0.67 ("most" sorted) and below 0.33 ("least" sorted) that belong to demographic groups in the ANES. 2020–2024 pools face-to-face and online samples.

Appendix B: Alternate measures of demographic sorting

This paper's demographic sorting measure relies on the same items from 1952 to 2024. This excludes social groups who were not included in early ANES waves, such as Asian Americans, the LGBT community, and born-again Christians. Therefore, the main analyses risk underestimating contemporary levels of demographic sorting by overlooking emerging social divisions. To investigate this, I construct another demographic sorting measure from 2008–2024 that includes more affiliations that were not measured in prior years (see Table OA.2). Nevertheless, I replicate Analyses #1–3 with this expanded measure.

Accounting for these categories raises demographic sorting levels by 0.01–0.02 points—a statistically significant but substantively modest increase[21]—and sorting's relationship with affective polarization does not change (see Figure OA.2). Furthermore, when using this expanded sorting measure as a substitute for the main measure from 2008–2024, I find that demographic sorting's explanatory power only increases from 2.2% to 3.4%. Thus, these divisions are not

[21]The increase is mainly driven by gun ownership, sexual orientation, and especially born-again Christian ID. From 2008–2024, these items' importance scores increase from 0.003 to 0.007, 0.001 to 0.006, and 0.009 to 0.019, respectively. The inclusion of the "not religious" category also appears to be impactful, increasing religious affiliation's 2024 importance score from 0.006 to 0.016.

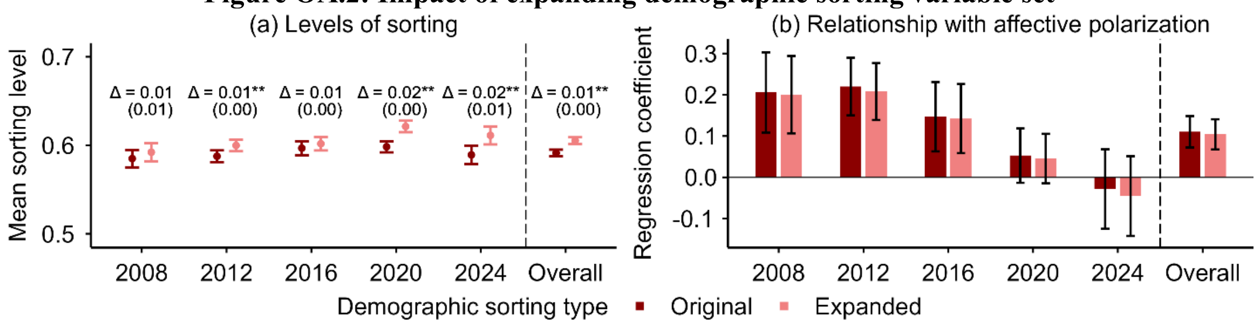
irrelevant, but they do not substantively alter the results. Emerging groups are either too small to affect overall estimates, or they are already covered to some degree by the original measure. For instance, secular and born-again respondents were likely represented at opposite ends of church attendance, and the original "Other" race category included AAPI and Native American voters.

Table OA.2: Items used in expanded demographic sorting measure

<u>Race</u>	<u>Religion</u>	Gender	Church attendance	**Born-again Christian**
• White • Hispanic	• Protestant • Jewish	Age	Marital status	**Sexual orientation**
• Black • Other	• Catholic • Other	Education	Census region	**Gun owner**
• Asian/Pacific Islander	**• Not religious**	Income	Union membership	**Parenthood**
• Native American		Urbanicity	Immigrant family	

Expanded demographic sorting for 2008–2024 is calculated with this set of predictors (new variables and categories in bold).

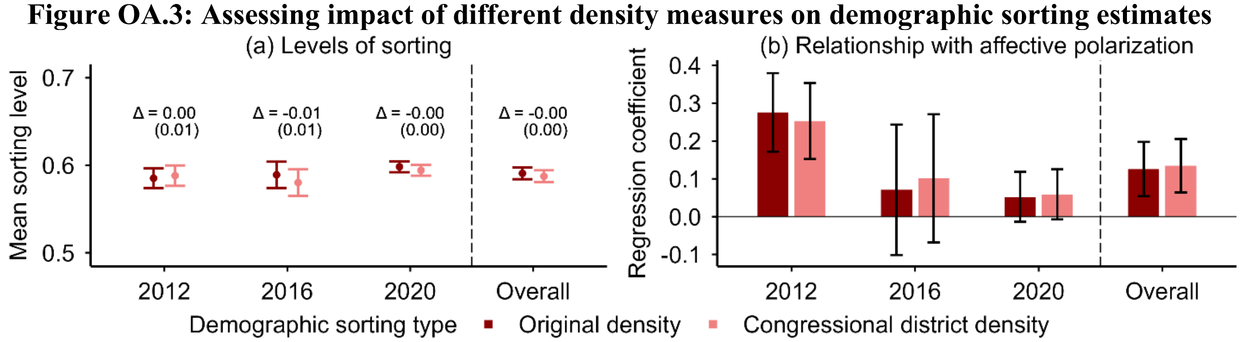
Figure OA.2: Impact of expanding demographic sorting variable set



Panel (a) compares mean demographic sorting estimates using original and expanded measures, with weighted two-sided t-tests for heuristic purposes (paired groups are not independent). Panel (b) compares these scores' relationships with affective polarization when controlling for attitudinal sorting (by attitude object) and survey mode. All results pool face-to-face and online samples to maximize statistical power. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$

In addition, the 2024 ANES does not measure urbanicity (and neither do the 2012 or 2016 online samples), so I construct three-point measures of urbanicity based on respondents' congressional district.[22] Mean demographic sorting estimates in 2012–2020 do not change when measuring urbanicity this way, and neither does sorting's relationship with polarization (see Figure OA.3). This indicates the substitution does not bias the 2024 results.

Figure OA.3: Assessing impact of different density measures on demographic sorting estimates



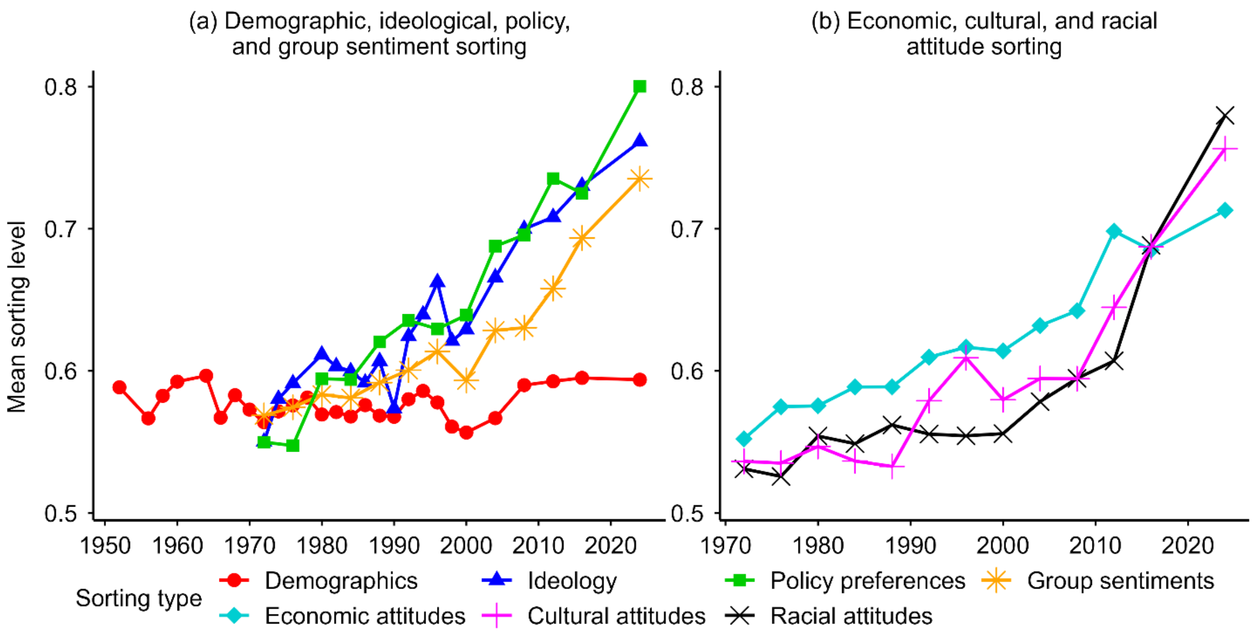
Panel (a) compares mean demographic sorting estimates using original and congressional-district measures, with weighted two-sided t-tests for heuristic purposes (paired groups are not independent). Panel (b) compares these scores' relationships with affective polarization when controlling for attitudinal sorting (by attitude object). 2012–2016 estimates use face-to-face interviews. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

[22]I use CityLab's Congressional Density Index for 2012–2020 (Montgomery 2019) and the Congressional District Health Dashboard's District Density Index for 2024 (NYU Langone Health 2022).

Appendix C: Excluding independents from sorting models

I measure partisan sorting by treating Republicans, Democrats, and (pure) Independents as three outcome classes and taking the difference between respondents' probabilities of in-party and out-party identification. Omitting independents from models does not alter my findings. (See also Figure OA.13 in Appendix I.)

Figure OA.4: Partisan sorting over time (independents excluded from sorting measure)



Mean partisan sorting scores in the ANES (face-to-face) each year, with independents excluded from predictive models. Demographic and ideological items are listed in Figure 2. Items selected for other measures are similar to those listed in Figures A1–A5 (see replication materials for full lists).

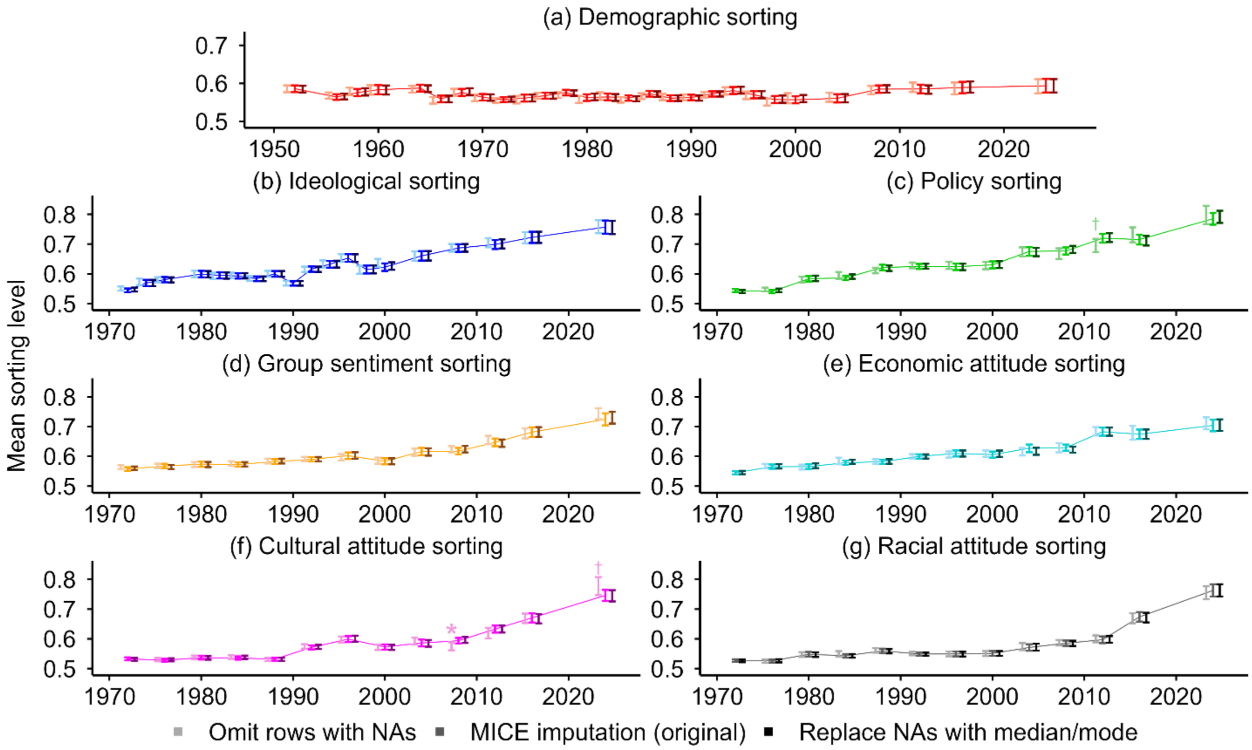
Appendix D: Assessing MICE imputation's effect on sorting

I use multivariate imputation by chained equations (MICE) to fill in missing values for each variable by iterating a series of estimates based on the item's relationship to other variables (van Buuren and Groothuis-Oudshoorn 2011). Each year, I run separate imputations for demographic, ideological, policy, and group sentiment variables, with each imputation undergoing ten iterations. I use predictive mean matching for continuous and ordinal variables, logistic regression for binary variables, and polytomous regression for nominal variables.

ANES data are low in missingness (mostly below 5% each year). The exceptions mainly involve missingness at random: policy items only asked of subsamples in 1972, 1984, and 2008, as well as urbanicity in 2000 when a large portion of the sample (44%) was interviewed over the telephone. In addition, refusal to disclose one's income was sometimes significantly higher than nonresponse rates for other demographic variables, with a maximum of 15.4% missingness in 2000. These low rates of missingness, combined with robust imputation, should minimize bias.

To confirm, I replicate Analysis #1 using two alternate approaches to handling missing data: listwise deletion of rows with missing values, and replacing NA values with medians (for ordinal and continuous variables) and modes (for nominal and binary variables). As I show in Figure OA.5, MICE imputation does not result in higher estimates of mean sorting—none of the 112 MICE-imputed estimates are even marginally significantly higher compared to both alternatives. This strongly indicates that imputation does not distort sorting levels.

Figure OA.5: Comparing mean sorting levels across different approaches to missing data

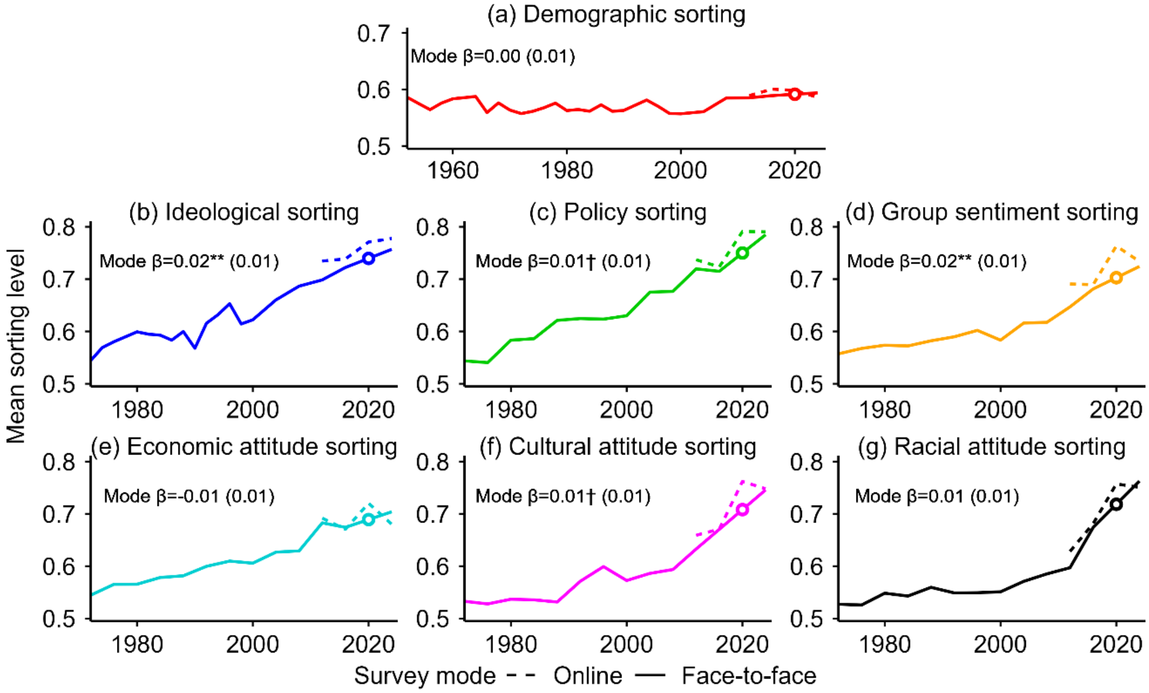


Mean partisan sorting scores in the ANES (face-to-face) each year, by missing-data approach. Confidence intervals on the left, center, and right display estimates based on omitting rows with NAs, MICE imputation (main text), and replacing NAs with medians and modes. No respondents received a full set of policy items in 1972, so five sorting types lack an "Omit rows with NAs" estimate that year. Significance markers come from weighted two-sided t-tests meant for heuristic purposes (paired groups are not independent). $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$

Appendix E: Online ANES time series trends

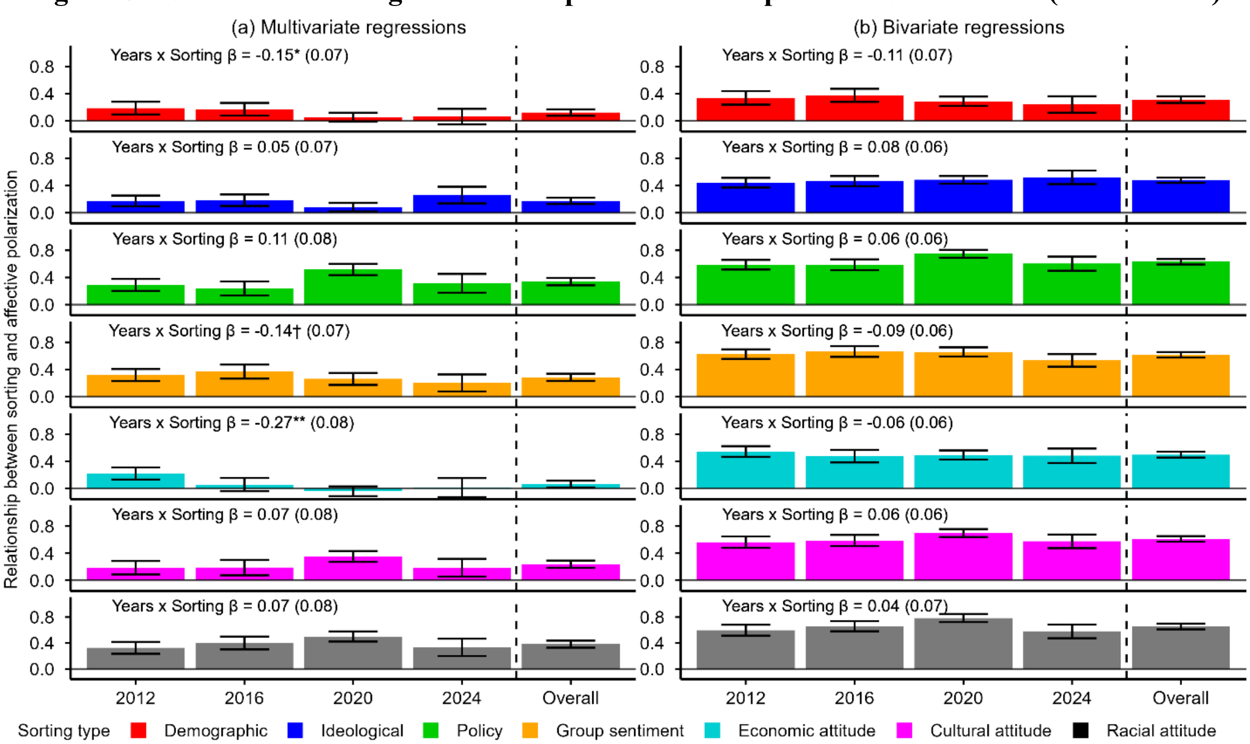
Online respondents express significantly higher levels of affective polarization than face-to-face (FTF) respondents (Thornton 2025; Tyler and Iyengar 2024; see also Figure 1). Below, I investigate whether demographic and attitudinal sorting show similar mode discrepancies. I calculate sorting levels for each year in the 2012–2024 online ANES time series (I run the online samples' MICE imputation and random forest models separately from the FTF data). I then pool the 2012, 2016, and 2024 FTF and online samples and regress each sorting type on time and mode dummies. I find similar time trends across modes, modest mode effects for some forms of attitudinal sorting, and no significant mode effect for demographic sorting (Figure OA.6). I also replicate Analysis #2 using only the online ANES time series (Figure OA.7). In this better-powered dataset (covering a much shorter time frame), demographic, economic attitude, and perhaps group sentiment sorting have become significantly less predictive of affective polarization since 2012.

Figure OA.6: ANES sorting trends over time, by survey mode



*Mean partisan sorting scores in the ANES each year, by survey mode. Face-to-face and online scores are estimated separately. 2020 has no face-to-face sample. Mode effects (†p < .10, *p < .05, **p < .01) are estimated from β_3 in $Sorting = \beta_0 + \beta_1 2016 + \beta_2 2024 + \beta_3 Mode + \varepsilon_i$.*

Figure OA.7: Partisan sorting's relationship with affective polarization over time (online series)



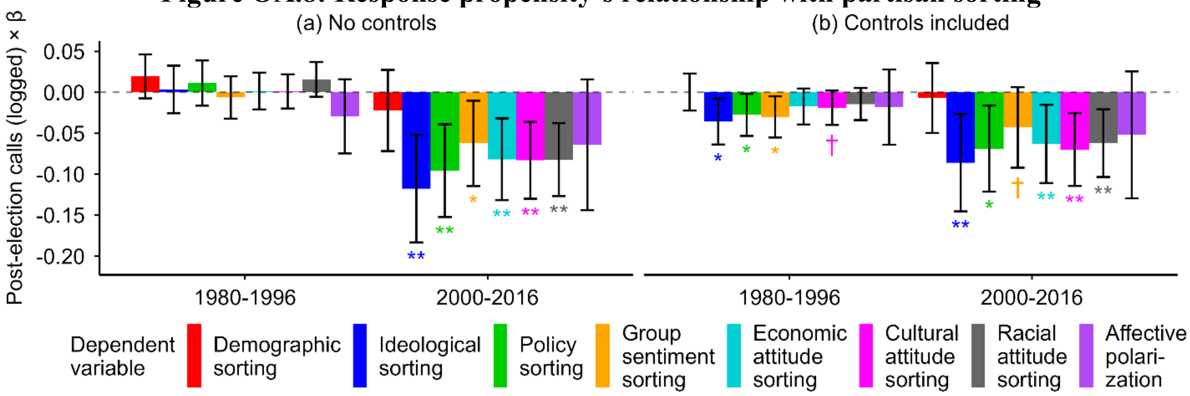
*Bars show regression coefficients from Equations #1 ("Overall") and #2 (but by year instead of decade). Years × Sorting β presents Equation #3's interaction coefficient. Affective polarization is the dependent variable in all models. Bivariate models use a single form of sorting, and Multivariate models use Demographic and Ideological sorting alongside either Policy and Group Sentiment sorting or Economic/Cultural/Racial Attitude sorting. †p < .10, *p < .05, **p < .01*

Appendix F: Nonresponse bias and rising attitudinal sorting

Public opinion scholarship is increasingly grappling with the implications of declining survey response rates and non-ignorable nonresponse (Bailey 2024; Clinton 2021; Tyler et al. 2025). Even high-quality datasets like the ANES risk overrepresenting politically engaged respondents, which could inflate attitudinal sorting increases over time (Cavari and Freedman 2018, 2023; cf. Mellon and Prosser 2021, 2025). I investigate this concern below across several analyses.

To begin, I show in Figure OA.8 that attitudinal sorting is significantly related to response propensity, as measured by the number of times interviewers contacted each respondent before conducting the post-election interview (Mellon and Prosser 2021; Peress 2010). This pattern emerges in the 2000s, with null or modest relationships beforehand. Demographic sorting is not related to response propensity, and neither (surprisingly) is affective polarization.

Figure OA.8: Response propensity's relationship with partisan sorting



This figure displays regression coefficients of the logged number of calls required to schedule each respondent's post-election interview, with sorting and polarization as dependent variables. Regressions cluster by interviewer; panel (b) regressions control for demographics and party ID. ANES 2024 post-election interview call data have not been released. †*p < .10,* **p < .05,* ***p < .01.*

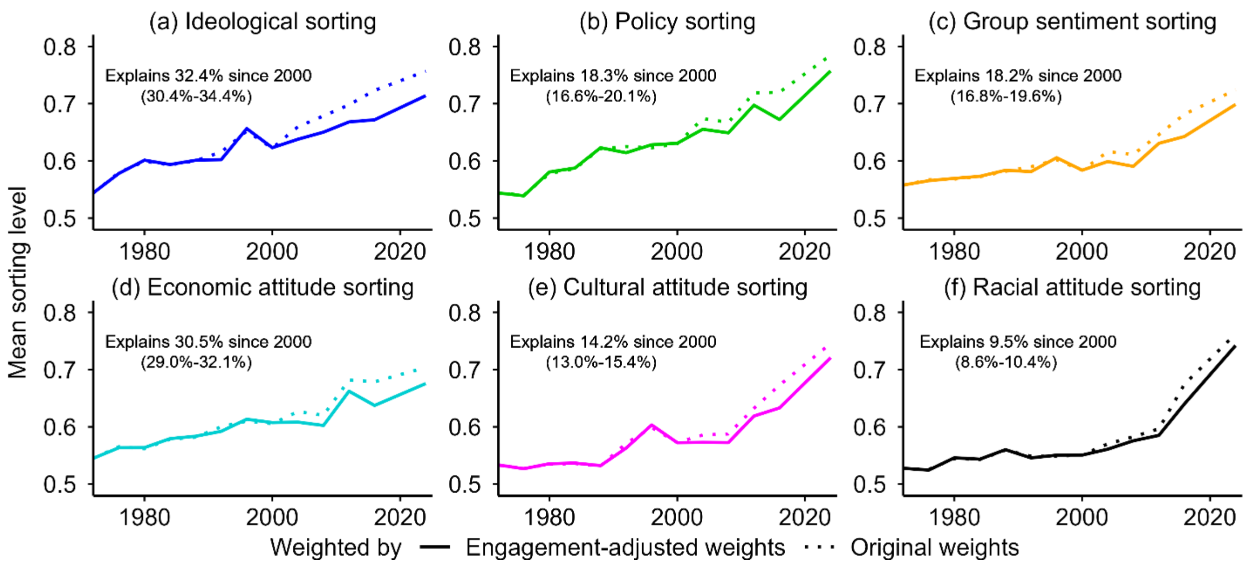
Although nonresponse bias is impossible to precisely correct for, I can provide an upper bound for its likely impact and benchmark it against observable confounders by drawing on two sensitivity analyses proposed by Hartman and Huang (2024). First, I treat political knowledge (measured by interviewer ratings) and interest as partially observed confounders and reweight the time series to match these variables' distributions in 1972–1980. Assuming *all* post-1980 increases in knowledge and interest only reflect nonresponse bias (likely a conservative assumption), accounting for this explains less than one third of any increase in attitudinal sorting since 2000 (see Figure OA.9). This suggests most of attitudinal sorting's rise is not an artifact of nonresponse.

Second, I benchmark a hypothetical unobserved confounder against the error of omitting various demographics from weights. In other words, I estimate how large an unobserved confounder would have to be relative to observable ones to explain a majority of attitudinal sorting's rise since 2000. To do so, I reweight the ANES, raking on age, education, gender, race, census region, and income targets based on 2024 Current Population Survey data. I then compute how much removing each target biases estimates of mean sorting levels by calculating:

- $\hat{\varepsilon}$: The difference between the full weight and the weight based on one omitted variable (i.e. the error caused by the omitted variable).

- $\widehat{R}^2_\varepsilon$: The weight error's variance (i.e. the variance in the difference between full weights and omitted-variable weights) divided by the full weight's variance. $\widehat{R}^2_\varepsilon$ is rescaled from 0 to 1, following Hartman and Huang (2024, Eq. 5).
- $\hat{\rho}_{\varepsilon,\text{Sorting}}$: The correlation between $\hat{\varepsilon}$ and attitudinal sorting.

Figure OA.9: Sorting trends over time, holding political interest and knowledge levels constant



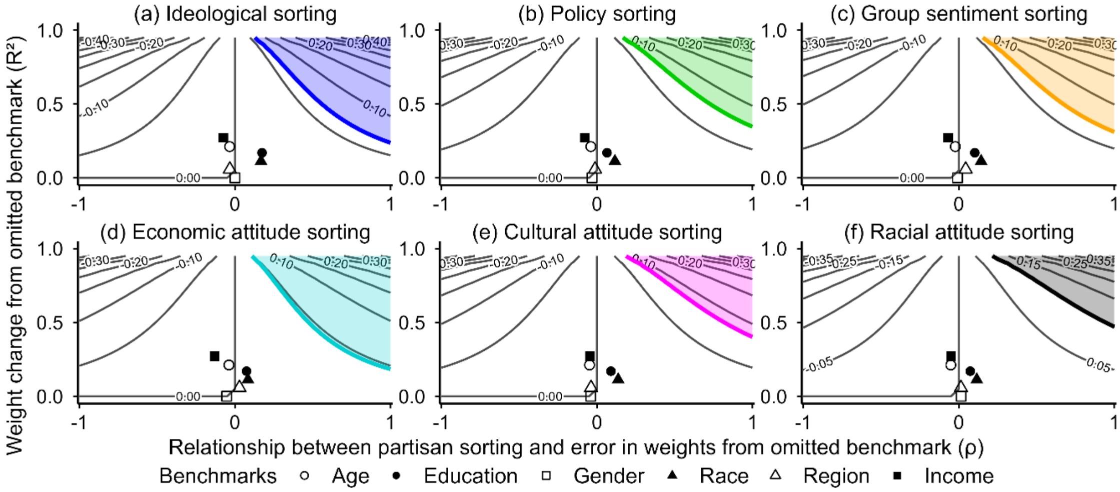
Mean sorting levels over time, when reweighting ANES (face-to-face) data to match political interest and knowledge levels in 1972–1980. Percentages display how much the main text's estimated increases in sorting from 2000–2024 decrease when reweighting (with 95% confidence intervals).

$\widehat{R}^2_\varepsilon$ represents how imbalanced the omitted demographic is: How much do the weights change when the demographic is removed? $\hat{\rho}_{\varepsilon,\text{Sorting}}$ shows how strongly related this imbalance is to partisan sorting. (How strongly does being misrepresented by the omitted-variable weights predict partisan sorting?) Both values must be high for a confounder to significantly bias estimated sorting levels. I therefore combine $\widehat{R}^2_\varepsilon$ and $\hat{\rho}_{\varepsilon,\text{Sorting}}$ to produce an overall estimate of bias from each demographic's omission, based on Hartman and Huang (2024, Theorem 3.1). I plot these estimates in Figure OA.10, alongside shaded "killer-confounder" regions that illustrate the bias necessary to explain at least half of the 2000–2024 increase.

I then compute each demographic's minimum relative confounding strength across all forms of sorting, which indicates how much stronger an unobserved confounder must be—relative to the benchmark demographic confounder—to bias estimates high enough to explain the majority of increased sorting from 2000–2024. For instance, if race has an MRCS of 2.0, unobserved confounders must inflate estimates by twice as much as not weighting by race.

As I show in Table OA.3, an unobserved confounder would need to exert at least 9.4 times the impact of omitting race, 7.1 times the impact of omitting education, or 5.9 times the impact of omitting income (though errors from omitting income go in the opposite direction). These are the minimum estimates across all forms of sorting—most are significantly higher. Thus, nonresponse bias, though a real concern, is unlikely to account for a majority of attitudinal sorting's rise.

Figure OA.10: Can omitted variable bias explain a majority of sorting's increases since 2000?



These bias contour plots' shaded regions represent the degree of bias from a confounder that would be necessary to explain a majority of the 2000–2024 increase in sorting. Demographic benchmarks' biases—represented by each point—are nowhere near these regions, so the effect of an unobserved confounder would have to be far larger (Table OA.3 lists specific estimates).

Table OA.3: How strong must omitted confounders be to explain a majority of increased sorting?

Demographic benchmark	$\hat{R}^2_\varepsilon$	Ideological sorting (Increase of 0.13 since 2000) $\hat{\rho}_{\varepsilon,\text{Sorting}}$	Bias	MRCS	Policy sorting (Increase of 0.16 since 2000) $\hat{\rho}_{\varepsilon,\text{Sorting}}$	Bias	MRCS	Group sentiment sorting (Increase of 0.14 since 2000) $\hat{\rho}_{\varepsilon,\text{Sorting}}$	Bias	MRCS
Age	0.21	−0.04	−0.00	−31.01	−0.04	−0.00	−36.02	−0.02	−0.00	−57.14
Education	0.17	0.17	0.01	7.09	0.07	0.00	24.84	0.10	0.01	14.48
Gender	0.00	0.00	−0.00	≤−100	−0.03	−0.00	≤−100	−0.01	0.00	≤−100
Race	0.11	0.17	0.01	9.37	0.12	0.00	17.68	0.14	0.01	13.15
Region	0.06	−0.03	−0.00	−67.89	−0.01	−0.00	≤−100	0.04	0.00	61.82
Income	0.27	−0.08	−0.01	−12.21	−0.08	−0.01	−15.33	−0.07	−0.00	−16.01

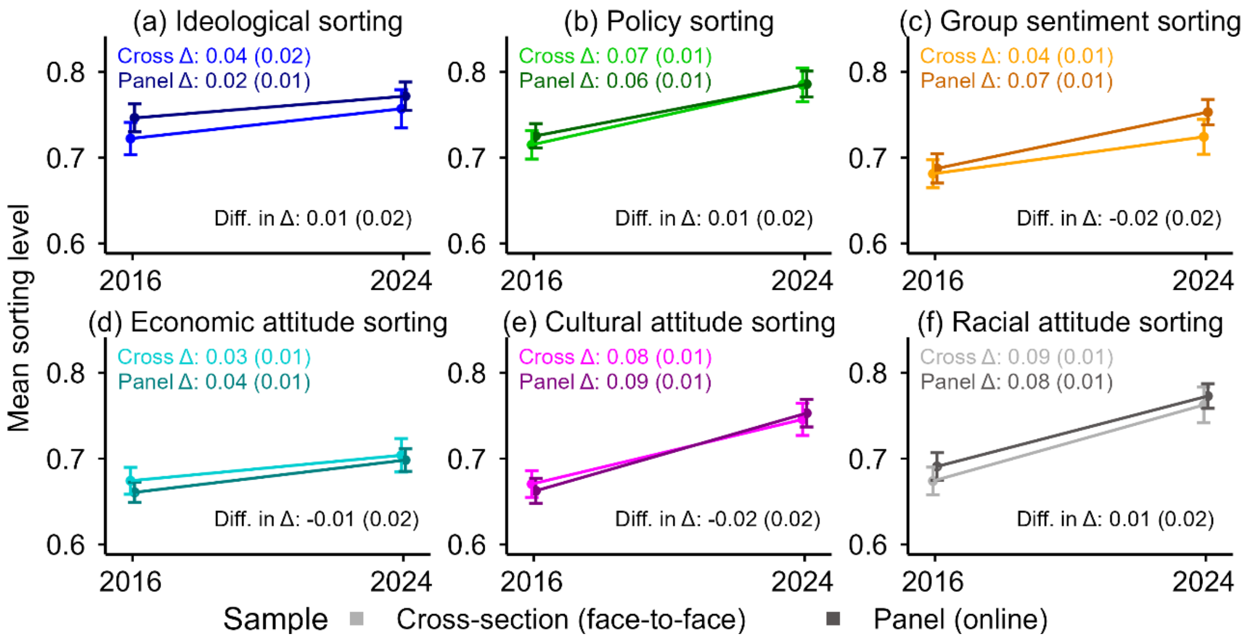
Demographic benchmark	$\hat{R}^2_\varepsilon$	Economic attitude sorting (Increase of 0.10 since 2000) $\hat{\rho}_{\varepsilon,\text{Sorting}}$	Bias	MRCS	Cultural attitude sorting (Increase of 0.17 since 2000) $\hat{\rho}_{\varepsilon,\text{Sorting}}$	Bias	MRCS	Racial attitude sorting (Increase of 0.21 since 2000) $\hat{\rho}_{\varepsilon,\text{Sorting}}$	Bias	MRCS
Age	0.21	−0.04	−0.00	−23.34	−0.05	−0.00	−33.32	−0.05	−0.00	−35.39
Education	0.17	0.07	0.00	14.12	0.09	0.01	20.09	0.07	0.00	28.42
Gender	0.00	−0.05	−0.00	≤−100	−0.04	−0.00	≤−100	0.01	0.00	≥100
Race	0.11	0.08	0.00	15.99	0.14	0.00	16.81	0.12	0.00	22.95
Region	0.06	0.03	0.00	67.15	−0.04	−0.00	−87.76	0.01	0.00	≥100
Income	0.27	−0.13	−0.01	−5.93	−0.05	−0.00	−29.77	−0.05	−0.00	−31.00

Benchmarks of how much each demographic can bias sorting estimates, based on 1) the degree of change in weights when omitting the demographic ($\hat{R}^2_\varepsilon$) and 2) the strength of sorting's relationship with the error in weights due to that omission ($\hat{\rho}_{\varepsilon,Sorting}$). Minimum relative confounding strength (MRCS) denotes how strong an unobserved confounder must be, relative to the benchmark, to explain a majority of sorting's increase since 2000. MRCS values > |100| are displayed as ≥ 100 or ≤ −100 to improve readability.

Finally, I compare attitudinal sorting changes in the 2016–2024 ANES panel to the cross-sectional results over the same period. Because the panel tracks the same subjects, over-time increases cannot be attributed to rising nonresponse bias. Using the same predictors to measure sorting each year (see Figures A1–A5), I calculate attitudinal sorting for the online panelists—removing 2016 face-to-face panel responses ensures increases are not due to changing modes (see Appendix E)—and compare their over-time changes to changes in the cross-sectional face-to-face time series. Although these samples are not perfectly comparable, given differences in survey mode and panelists' higher response propensity, they enable a useful test: Do attitudinal sorting increases persist when they cannot be attributed to changes in sample composition?

I find they do persist: Attitudinal sorting increases in the panel sample are not significantly smaller than cross-sectional trends (Figure OA.11). Thus, attitudinal sorting increases are reproduced when tracking the same respondents over time, which indicates even the large jumps in sorting from 2016–2024 are not primarily artifacts of declining response rates.

Figure OA.11: Changes in attitudinal sorting, 2016–2024 (cross-section versus panel)



A comparison of partisan sorting's 2016–2024 increases in the ANES's face-to-face cross-sections ("Cross Δ") and online panel ("Panel Δ"). "Diff. in Δ" is the difference between those differences. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

Appendix G: The influence of early demographic sorting

Demographics are powerful early forces in the "funnel of causality" that shapes voters' social identities, partisanship, voting behavior, and policy preferences (see Campbell et al. 1960). Given this, even though cross-sectional evidence indicates demographic sorting cannot explain rising affective polarization in the electorate (even indirectly), it remains useful to test whether early demographic sorting predicts affective polarization and attitudinal sorting later in life.

I investigate this using the 1965–1997 Youth-Parent Socialization Panel, which tracks a cohort of 935 voters from ages 18–50. I apply the same predictive modelling approach to measure demographic sorting in 1965 and 1973 and attitudinal sorting across all four waves. Because the 1965 wave includes a sparse set of attitudes, I construct a single "omnibus" measure of attitudinal sorting that combines policy preferences and group sentiments (see Table OA.4 for items used to calculate demographic sorting and 1965 omnibus attitudinal sorting). I also create an omnibus measure of 1973 attitudinal sorting—based on all items listed in Figure OA.16—for use as an independent variable for the regressions shown in Table OA.5. I then divide those items into the usual six measures of attitudinal sorting for the 1973–1997 waves. These serve as dependent variables in Tables OA.6–OA.7 and both independent and dependent variables in Table OA.8.

Table OA.4: Predictors used to calculate partisan sorting in Youth-Parent Socialization Study

1965 Demographics		1973 Demographics		1965 Attitudes	
Race	Church attendance	Race	Census region	School prayer	Protestants FT
Gender	Parent education	Gender	Urbanicity	School integration	Jews FT
Religion	**Parent income**	Religion	Church attendance	Labor unions FT	Whites FT
Census region	**Parent union member**	Education	Union membership	Big business FT	Blacks FT
Urbanicity	**Parent marital status**	Income	Marital status	Catholics FT	Southerners FT

This table lists the predictors used to calculate demographic sorting in 1965 and 1973, as well as "omnibus" attitudinal sorting in 1965. See Appendix L for predictors used to calculate attitudinal sorting in 1973, 1982, and 1997. Bolded items were obtained from parent interviews and matched to youth panelists. FT=feeling thermometers.

First, I test whether demographic sorting at ages 18 and 26 predicts later affective polarization. Table OA.5 suggests demographic sorting in 1965 predicts future polarization in 1973 and perhaps even 1997, but the latter effect fades when controlling for attitudinal sorting and is not reproduced in 1982. Demographic sorting in 1973 does not significantly predict any future affective polarization. Thus, I find only sporadic evidence that early-life demographic sorting can influence future affective polarization. The absence of strong evidence here is telling, as the models do not even control for lagged affective polarization (which was not measured in 1965).

Table OA.5: Does early sorting predict future affective polarization?

	1965 sorting as independent variable(s)						1973 sorting as independent variable(s)			
	FT Pol. 1973	FT Pol. 1973	FT Pol. 1982	FT Pol. 1982	FT Pol. 1997	FT Pol. 1997	FT Pol. 1982	FT Pol. 1982	FT Pol. 1997	FT Pol. 1997
Demographic sorting	0.11^{*} (0.05)	$0.09^{\dagger}$ (0.05)	0.08 (0.05)	0.08 (0.06)	0.09^{*} (0.05)	0.06 (0.05)	0.07 (0.07)	0.04 (0.07)	0.09 (0.06)	0.07 (0.06)
Attitudinal sorting		0.06 (0.06)		0.02 (0.08)		0.10 (0.07)		0.19^{**} (0.05)		0.08 (0.06)
Intercept	0.06^{*} (0.03)	0.04 (0.03)	0.14^{**} (0.03)	0.14^{**} (0.03)	0.17^{**} (0.03)	0.13^{**} (0.04)	0.15^{**} (0.04)	0.06 (0.05)	0.17^{**} (0.03)	0.13^{**} (0.04)
N	604	604	708	708	766	766	687	687	720	720

This table reports the results of regressing future affective polarization ("FT Pol.") on 1965–1973 demographic sorting and omnibus attitudinal sorting in the Youth-Parent Socialization Study. Omnibus attitudinal sorting pools all attitudes in Table OA.4 and Figure OA.16 for 1965 and 1973 measures, respectively. Robust standard errors in parentheses. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

Second, I test whether early demographic sorting can indirectly increase affective polarization by shaping future attitudinal sorting. In Tables OA.6–OA.7, I regress all six 1973–1997 attitudinal sorting scores on 1965 demographic and attitudinal sorting. 1965 demographic sorting at least marginally predicts higher sorting across three sets of attitudes in 1973, but none by 1997. (Notably, 1965 attitudinal sorting has even less predictive power, but this can likely be attributed to the paucity of available items.) Table OA.8 replicates this analysis with 1973's sorting measures: Demographic sorting at age 26 significantly predicts four forms of attitudinal sorting at age 35 and only one form—racial attitude sorting—by age 50.

Overall, early-life demographic sorting may be marginally predictive of future affective polarization and can sometimes predict later attitudinal sorting, but few relationships are consistent across specifications or durable in the long run. These findings—based on one cohort born near the start of the ANES—are far from definitive. But they are consistent with Analysis #2, in that they do not support a causal account in which demographic sorting spurred the rise in affective polarization, even indirectly.

Table OA.6: Does 1965 sorting predict future ideological, policy, and group sentiment sorting?

	Ideological sorting			Policy sorting			Group sentiment sorting		
	1973 sorting	1982 sorting	1997 sorting	1973 sorting	1982 sorting	1997 sorting	1973 sorting	1982 sorting	1997 sorting
Demographic sorting (1965)	0.03 (0.05)	0.05 (0.06)	−0.00 (0.06)	0.06† (0.03)	0.05 (0.04)	−0.05 (0.03)	0.07† (0.04)	0.08 (0.05)	0.07 (0.04)
Attitudinal sorting (1965)	0.06 (0.05)	−0.02 (0.08)	−0.05 (0.08)	0.08† (0.04)	0.04 (0.04)	0.07† (0.04)	0.03 (0.04)	0.04 (0.05)	−0.06 (0.06)
Intercept	0.53** (0.04)	0.60** (0.04)	0.71** (0.04)	0.47** (0.03)	0.51** (0.03)	0.56** (0.03)	0.49** (0.02)	0.52** (0.03)	0.61** (0.03)
N	684	711	767	684	711	767	684	711	767

This table reports the results of regressing 1973–1997 ideological, policy, and group sentiment sorting on 1965 demographic and (omnibus) attitudinal sorting in the Youth-Parent Socialization Study. Table OA.4 lists items used to calculate 1965 sorting; Figure OA.16 lists items used to calculate 1973–1997 sorting. Robust standard errors in parentheses. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

Table OA.7: Does 1965 sorting predict future economic, cultural, and racial attitude sorting?

	Economic attitude sorting			Cultural attitude sorting			Racial attitude sorting		
	1973 sorting	1982 sorting	1997 sorting	1973 sorting	1982 sorting	1997 sorting	1973 sorting	1982 sorting	1997 sorting
Demographic sorting (1965)	0.06 (0.04)	0.08 (0.05)	0.06 (0.05)	0.02 (0.03)	0.03 (0.03)	−0.02 (0.04)	0.07* (0.03)	0.10* (0.04)	0.06 (0.04)
Attitudinal sorting (1965)	0.00 (0.05)	−0.02 (0.05)	−0.02 (0.06)	0.03 (0.03)	0.04 (0.04)	−0.09 (0.06)	0.03 (0.03)	0.04 (0.04)	0.06 (0.05)
Intercept	0.52** (0.03)	0.56** (0.03)	0.56** (0.04)	0.51** (0.03)	0.50** (0.02)	0.66** (0.03)	0.48** (0.02)	0.47** (0.03)	0.48** (0.03)
N	684	711	767	684	711	767	684	711	767

This table reports the results of regressing 1973–1997 economic, cultural, and racial attitude sorting on 1965 demographic and (omnibus) attitudinal sorting in the Youth-Parent Socialization Study. Table OA.4 lists items used to calculate 1965 sorting; Figure OA.16 lists items used to calculate 1973–1997 sorting. Robust standard errors in parentheses. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

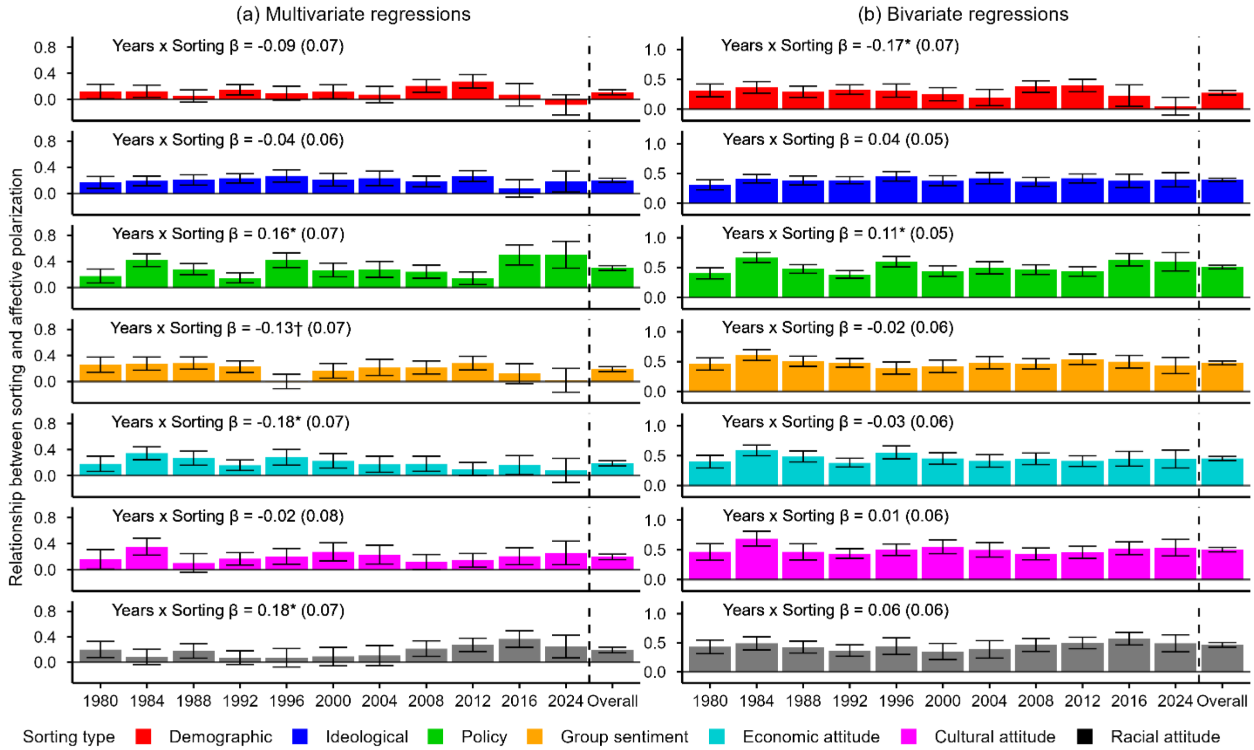
Table OA.8: Does 1973 sorting predict future attitudinal sorting?

	Ideological sorting	Policy sorting	Group sentiment sorting	Economic attitude sorting	Cultural attitude sorting	Racial attitude sorting
Dependent variable: 1982 attitudinal sorting						
Demographic sorting (1973)	0.03 (0.05)	0.09** (0.04)	0.12** (0.05)	0.12** (0.05)	0.03 (0.03)	0.15** (0.04)
Attitudinal sorting (1973)	0.34** (0.05)	0.32** (0.04)	0.26** (0.05)	0.27** (0.05)	0.30** (0.04)	0.30** (0.04)
Intercept	0.40** (0.04)	0.34** (0.03)	0.37** (0.03)	0.38** (0.04)	0.37** (0.02)	0.31** (0.03)
Dependent variable: 1997 attitudinal sorting						
Demographic sorting (1973)	−0.04 (0.06)	0.06 (0.04)	0.08 (0.05)	0.07 (0.05)	−0.05 (0.05)	0.19** (0.03)
Attitudinal sorting (1973)	0.26** (0.06)	0.21** (0.04)	0.24** (0.06)	0.17** (0.05)	0.34** (0.05)	0.13** (0.05)
Intercept	0.56** (0.05)	0.43** (0.03)	0.45** (0.04)	0.46** (0.03)	0.45** (0.04)	0.38** (0.03)

This table reports the results of regressing each form of attitudinal sorting in 1982–1997 on 1973 demographic sorting and the lagged 1973 attitudinal sorting measure that matches the dependent variable. Figure OA.16 lists items used to calculate 1973–1997 sorting. Robust standard errors in parentheses. N_{1982}=688; N_{1997}=721. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

Appendix H: Binning by year in Analysis #2 regressions

Figure OA.12: Partisan sorting's relationship with affective polarization over time (yearly)



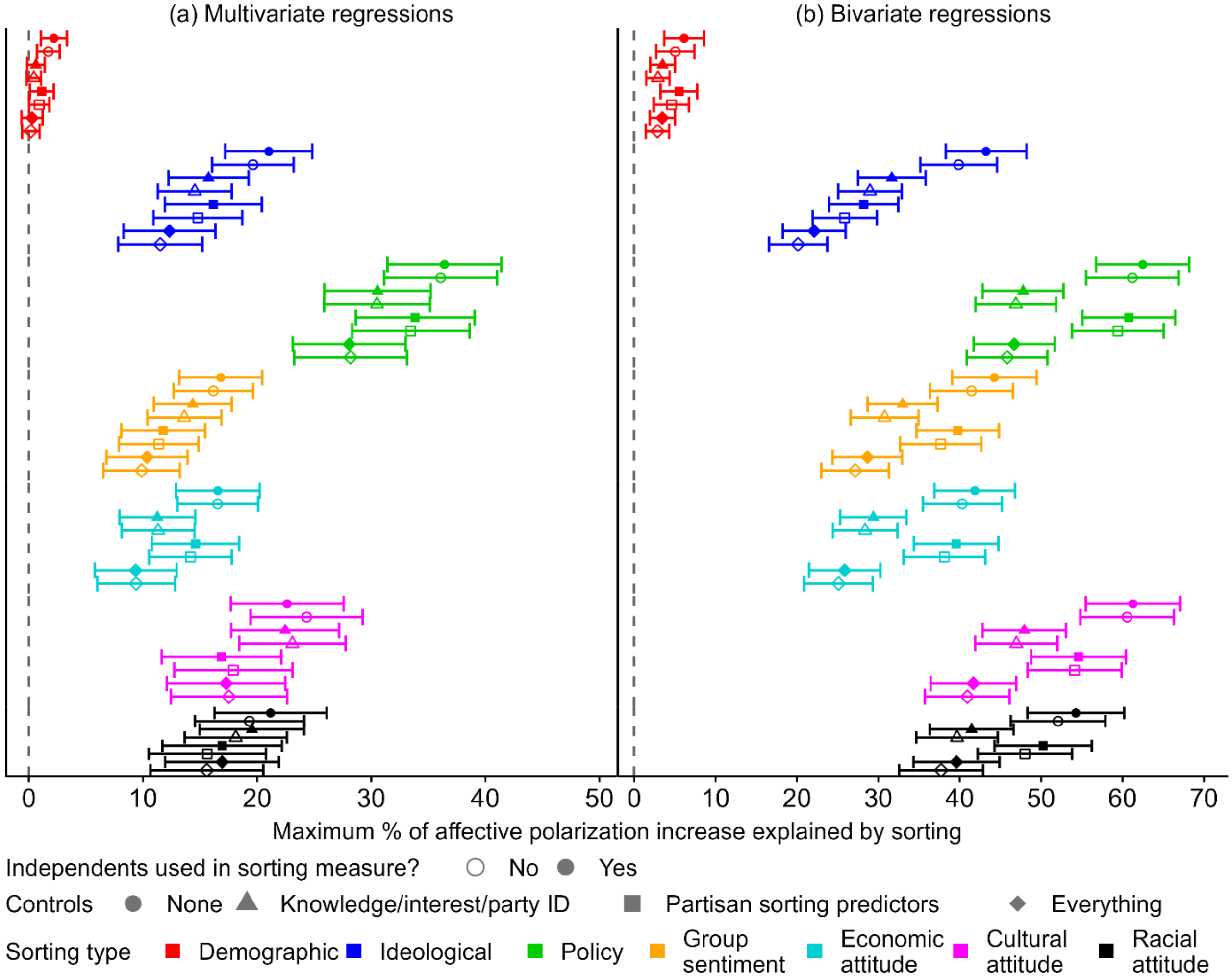
Bars show regression coefficients from Equations #1 ("Overall") and #2 (but by year instead of decade). Years × Sorting β presents Equation #3's interaction coefficients. Affective polarization is the dependent variable in all models. Bivariate models use a single form of sorting, and Multivariate models use Demographic and Ideological sorting alongside either Policy and Group Sentiment sorting or Economic/Cultural/Racial Attitude sorting. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

Appendix I: Decompositions across several specifications

I investigate how the results of Analysis #3 change when including three sets of controls: political knowledge, political interest, and party ID; the predictors used to calculate partisan sorting; and both sets combined. By including predictors as controls, I can assess whether partisan sorting's predictive power is truly a consequence of declining cross-pressures, or if it can simply be attributed to traits' "direct effects" on affective polarization. For instance, if higher education is predictive of increased polarization, irrespective of its alignment with party, then this would be a "direct effect" of the trait which the sorting measure indirectly captures. Similarly, if policy extremity is predictive of higher affective polarization even when it does not align with party, this is a direct effect that could be conflated with the influence of partisan sorting (see Horan 1971).

As Figure OA.13 shows, adding predictor controls modestly reduces the relationship between partisan sorting and polarization in some multivariate models—likely due to the direct use of feeling thermometers in both sides of the regression—but attitudinal sorting continues to have substantial explanatory power, while demographic sorting's estimate goes from 2.2% to 1.1%. Excluding independents from sorting predictive models (see Appendix C) makes no difference.

Figure OA.13: How much could each type of sorting have contributed to rising polarization?



Results obtained by applying Equations #4–6 in Analysis #3 across 16 specifications, for each form of sorting. Predictor controls are the demographic and ideological items listed in Figure 2, as well as all predictors consistently available since 1980: support for abortion, aid to Blacks, defense spending, and a government jobs program; and feeling thermometers for big business, labor unions, Blacks, whites, Hispanics, and either feminists or (in 1980–1988 and 1996) the women's movement. To account for nonlinear relationships, attitudes are as ordered factors, and each numeric feeling thermometer predictor is paired with a quadratic term. Multivariate models include all predictors; bivariate models include only the predictors for that sorting measure.

Appendix J: Results when using same attitudes over time

Using the most predictive items each year allows for analyses going back to 1972 and accounts for issues that have faded from public debate (e.g., racial busing, Vietnam) or recently emerged (e.g., border wall, transgender policy). However, this "top-item" approach risks conflating over-time improvements in measurement with genuine sorting. Given this, I construct two types of sorting measures that hold items constant from 1992–2024.

First, the "same-item" measures use consistent sets of predictors to measure sorting in every election year since 1992 (see Table OA.9). However, there is a shortage of group sentiment and racial attitude items consistently available from 1992–2024. Many available items—such as the racial group feeling thermometers—are also among the weaker predictors used by the original measures (see Figure A5). This presents the opposite problem: Just as sorting trends can be inflated when the quality of items increases over time, they can be deflated by a dearth of available items.

Fortunately, some variables were excluded because they were omitted from just one or two waves (e.g. illegal immigrants feeling thermometer). Given this, I can construct a second measure based on a pool of all items that are at least available in 1992 and 2024 (see bolded items in Table OA.9), with random forests selecting the pool's most predictive items each year. This second "same-pool" measure partially corrects for sparse item sets when estimating over-time change, though even this approach cannot capture emerging (and fading) issue divides.

Table OA.9: Items used to calculate "same-item" and "same-pool" sorting scores (1992–2024)

Sorting	Variables
Policy	Affirmative action in hiring, Aid to Blacks, Death penalty, Defense spending, Emphasize traditional values, **Gay adoption**, Government health insurance, Government jobs program, Government services vs. spending, Increase or decrease immigrants, Isolationism, **Less or more government**, Place limits on imports, Protect gays from discrimination, Should adjust moral views, Spending on assistance to the poor, Spending on crime, **Spending on environment**, Spending on public schools, Spending on social security, Spending on welfare, **Urban unrest**, When should abortion be allowed
Group sentiment	**Asian Americans FT**, Big business FT, Blacks FT, Christian fundamentalists FT, Feminists FT†, Gays and lesbians FT, Hispanics FT, **Illegal immigrants FT**, **Jews FT**, Labor unions FT, **Police FT**, Whites FT
Economic attitude	Big business FT, Government health insurance, Government jobs program, Government services vs. spending, Labor unions FT, **Less or more government**, Place limits on imports, Spending on assistance to the poor, Spending on public schools, Spending on social security, Spending on welfare
Cultural attitude	Christian fundamentalists FT, Death penalty, Emphasize traditional values, Feminists FT†, **Gay adoption**, Gays and lesbians FT, **Jews FT**, **Police FT**, Protect gays from discrimination, Should adjust moral views, Spending on crime, **Spending on environment**, **Urban unrest**, When should abortion be allowed
Racial attitude	Affirmative action in hiring, Aid to Blacks, **Asian Americans FT**, Blacks FT, Hispanics FT, **Illegal immigrants FT**, Increase or decrease immigrants, Whites FT

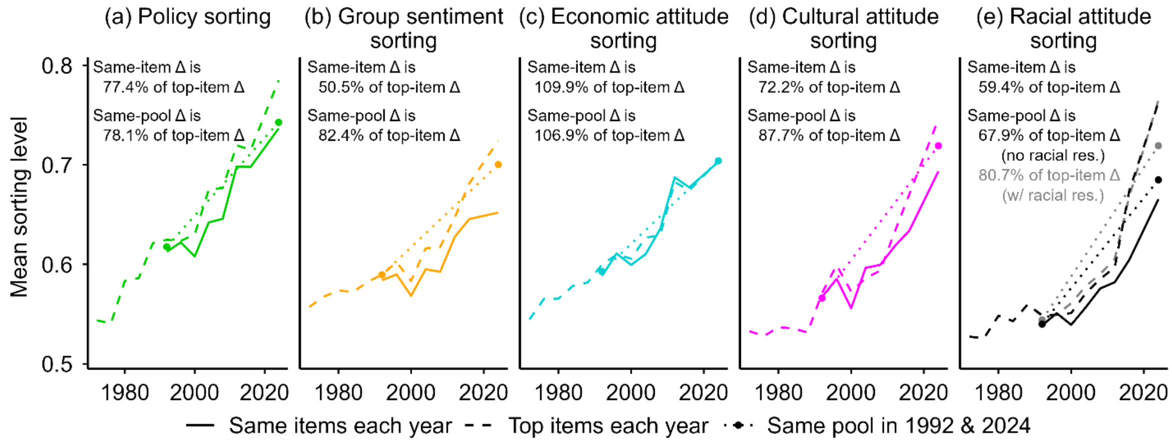
FT=feeling thermometers. All items are available predictors for "same-pool" sorting (calculated in 1992 and 2024 only). Non-bolded items are also used for "same-item" sorting measures throughout 1992–2024. I cap the number of same-pool predictors to 14 for policy and group sentiment sorting, and to eight for economic, cultural, and racial attitude sorting (matching the main text).
†In 1996, I substitute the women's movement FT for the feminists FT; the items correlate at r = 0.638 in 1992 and 2000 (pooled).

When using the same items every year, I replicate over 70% of policy, economic attitude, and cultural attitude sorting's increases (Figure OA.14). The same-item measures also reproduce a modest majority of group sentiment and racial attitude sorting's increases. The same-pool trends close much of this deficit for group sentiment sorting, but not racial attitude sorting. However, Figure OA.14 also displays racial attitude sorting trends when adding racial resentment items as potential predictors for both the top-item and same-pool measures.[23] Including these items accounts for a larger portion of the deficit, suggesting that item scarcity can explain much of the smaller same-item increases in the full 1992–2024 time series.

In addition, Figure OA.15 replicates Analysis #3 for 1992–2024. Because the estimates are based on same-item measures, there are smaller estimated contributions (particularly for group sentiment sorting), but the overall patterns are roughly the same as those shown by Figure OA.13.

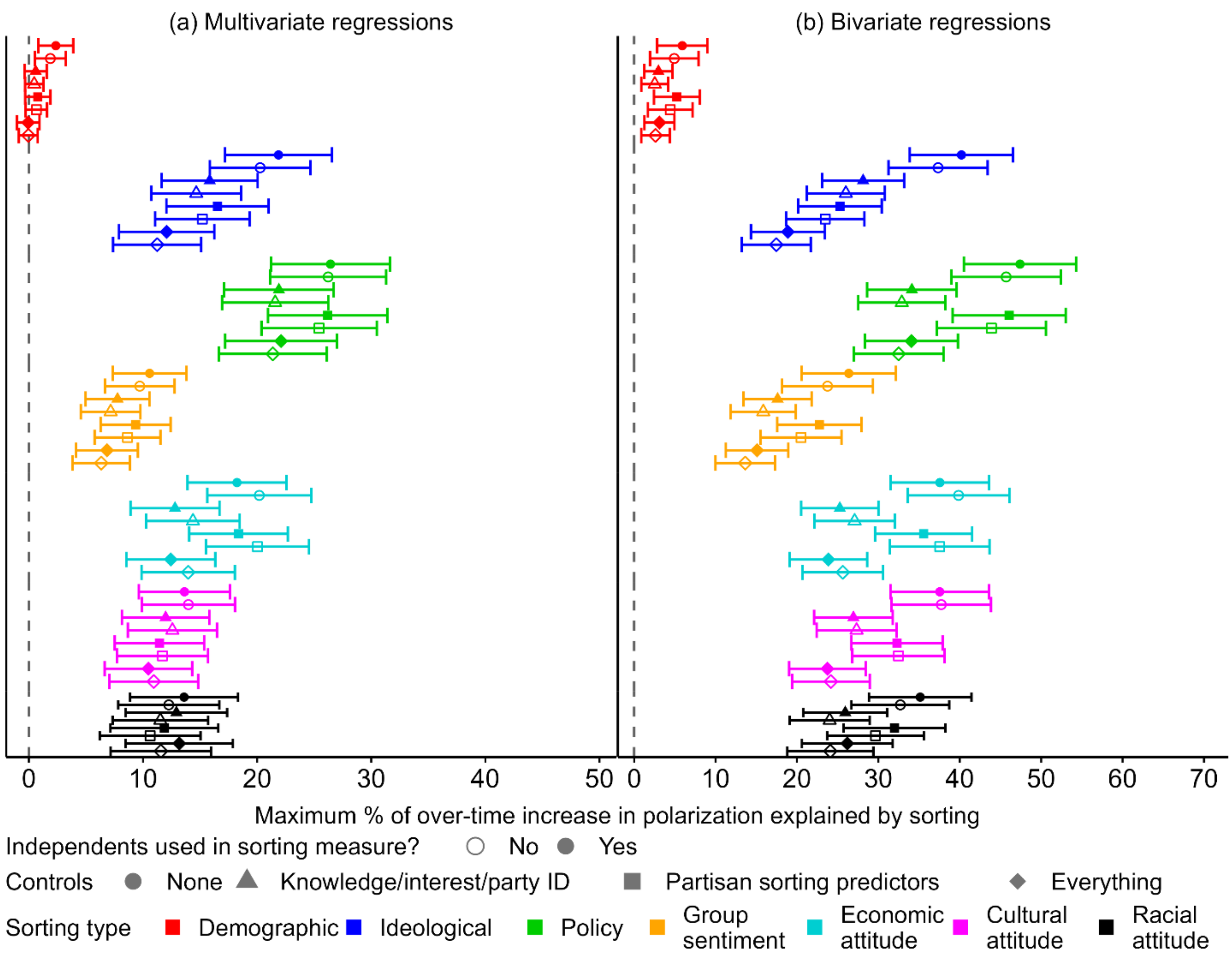
[23] I exclude racial resentment from the main measures due to uncertainty around whether the items should be counted as policy preferences or group sentiments. As Figure OA.14e's gray dashed line shows, their exclusion does not affect over-time trends.

Figure OA.14: How do sorting levels differ when using the same items over time?



Increases in mean sorting levels from 1992–2024 when using the same items every year ("Same-item Δ") and the same pool of predictors in 1992 and 2024 ("Same-pool Δ"), compared to sorting increases shown in the main text, which are based on the most predictive items each year ("Top-item Δ"). See Table OA.9 for items used to calculate same-item and -pool sorting.

Figure OA.15: How much could sorting have contributed to rising polarization? (same predictors)



Replications of Analysis #3 using the same items to measure sorting every year (from 1992–2024). See Appendix I for details. "Partisan sorting predictor" controls include all predictors used, not just the subset available from 1980–2024 (as in Appendix I). Items used to calculate sorting can be found in Figure 2 and Table OA.9.

Appendix K: Analysis #4 replication details

Mason (2018) and Mason and Wronski's (2018) sociopartisan sorting scales—replicated in Analysis #4—code respondents' group affiliations by their alignment with party ID, weight each affiliation by identity strength, and average all identities together (see Table OA.10 for formulas).

Table OA.10: Sociopartisan sorting formulas used in Analysis #4

Table 4 2011 YouGov (Mason 2018, Table A.5)	Table 5 (Columns #1–3) 2016 CES Module (Mason and Wronski 2018, Table A.3)	Table 5 (Columns #4–9) 2012 ANES; UCLA Nationscape (waves #13, 22, 32, 40, 50, 75, & 76)
	Step 1: Alignment scores	
Ideology Conservative Republican = 1 Liberal Democrat = 1 Liberal Republican = −1 Conservative Democrat = −1 Moderate = 0 **Tea Party & Evangelical** Republican = 1 Democrat = −1 Not-affiliated with group = NA **Secular & Black** Democrat = 1 Republican = −1 Not-affiliated with group = NA	**Ideology** Conservative Republican = 1 Liberal Democrat = 1 Liberal Republican = −1 Conservative Democrat = −1 Moderate = 0 **White & Christian** Republican = 1 Democrat = −1 Not-affiliated = NA **Black, Hispanic, & Secular** Democrat = 1 Republican = −1 Not-affiliated with group = NA	**Ideology** Conservative Republican = 1 Liberal Democrat = 1 Liberal Republican = −1 Conservative Democrat = −1 Moderate = 0 **Race** White Republican = 1 Non-white Democrat = 1 White Democrat = −1 Non-white Republican = −1 **Religion** Christian Republican = 1 Non-Christian Democrat = 1 Christian Democrat = −1 Non-Christian Republican = −1
	Step 2: Identity scores	
Party Partisan identity scale **Ideology, Tea Party, Evangelical, Secular, & Black** Identity scale × matching Alignment	**Party** Folded 7-point party ID **Ideology** Folded 7-point ideology × Alignment **White, Black, Hispanic, Christian, & Secular** Group Closeness rating × matching Alignment	**Party** Partisan identity importance **Ideology** Folded 7-point ideology × Alignment **Race & Religion** Identity Importance rating × matching Alignment
	Step 3: Calculate final scores	
Sociopartisan sorting Mean of all identity scores **Political identity sorting** Mean of Party, Ideology, & Tea Party identity scores **Demographic identity sorting** Mean of Evangelical, Secular, & Black identity scores **Individual scores** Rescale from 0−1, replace NAs with 0.5[†]	**Sociopartisan sorting** Mean of all identity scores **Political identity sorting** Mean of Party and Ideology identity scores **Demographic identity sorting** Mean of White, Black, Hispanic, Christian, & Secular identity scores **Individual scores** Rescale from 0−1, replace NAs with 0.5[†]	**Sociopartisan sorting** Mean of all identity scores **Political identity sorting** Mean of Party and Ideology identity scores **Demographic identity sorting** Mean of Race and Religion identity scores **Individual scores** Rescale from 0−1

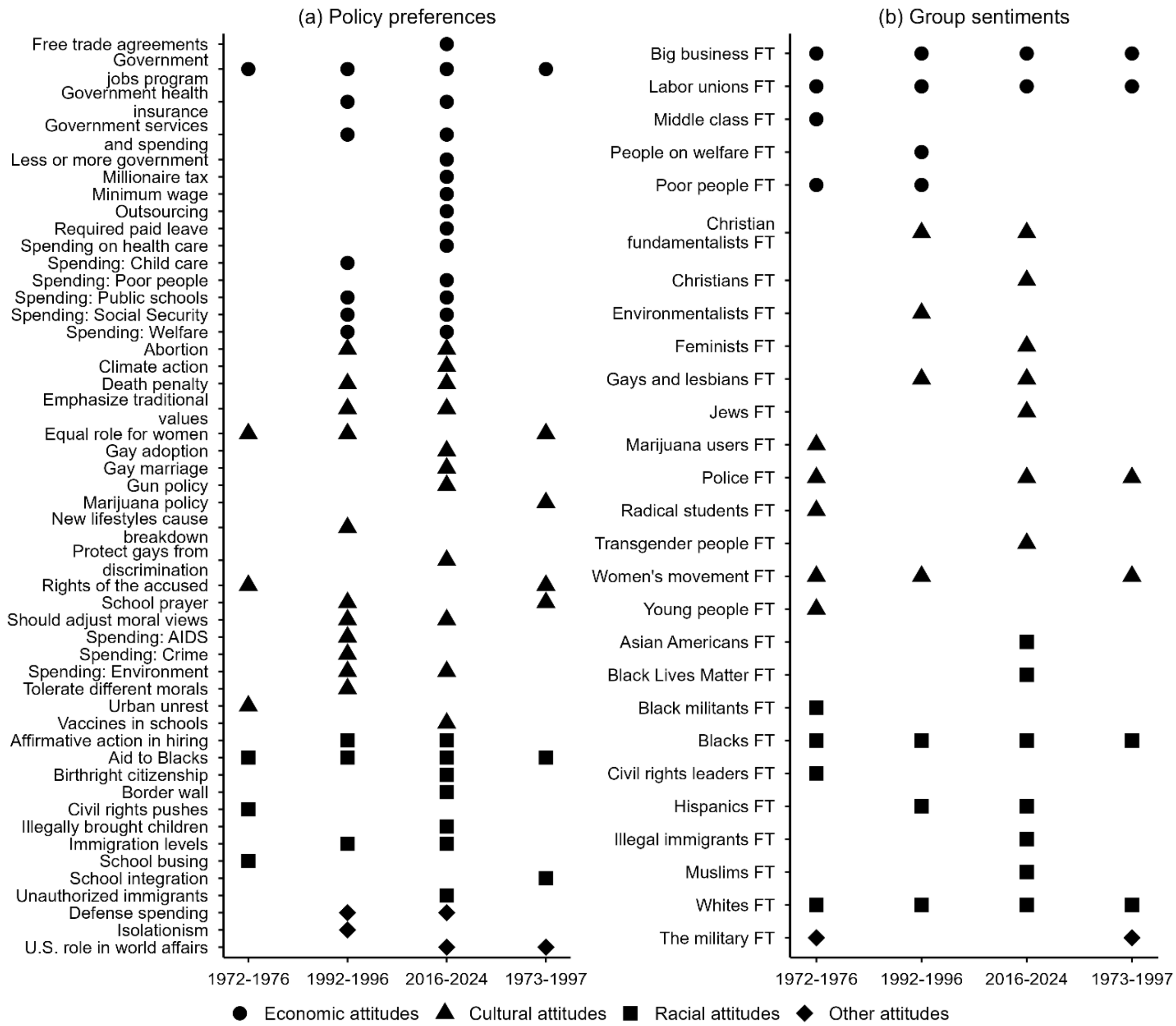
This table describes the calculations for all measures of identity sorting in Analysis #4, based on Mason (2018) and Mason and Wronski (2018). [†]NAs are replaced with 0.5 so regressions shown in Table 4 (columns #4, #6, #8, and #10) have complete cases; no respondents have non-NA values for all individual identities. I carry this to Table 5, column #3, but it is unnecessary for #6 or #9 because the 2012 ANES and Nationscape both measure identity strength for respondents' overall racial and religious affiliations (rather than specific identities, such as Black and Evangelical).

Table 4 uses Mason's (2018) measures of issue extremity and constraint. Extremity is an additive scale of five policy items (immigration, health care, abortion, gay marriage, and deficit reduction), all folded at the midpoint and weighted by issue importance. Constraint is the absolute value of the difference between the proportions of conservative and liberal policy positions.

Finally, Mason (2018) miscodes three demographic controls used in Table 4 (raw education and income variables are used, and age is improperly rescaled and binned). I only use these miscoded variables in column #1's exact reproduction; the rest of Table 4's regressions use corrected versions. Mason (2018) and Mason and Wronski (2018) also do not use survey weights. I exclude weights from the exact reproduction in Table 4 (column #1) but otherwise use them.

Appendix L: Items used to calculate sorting in panels

Figure OA.16: Items used to calculate partisan sorting in panels



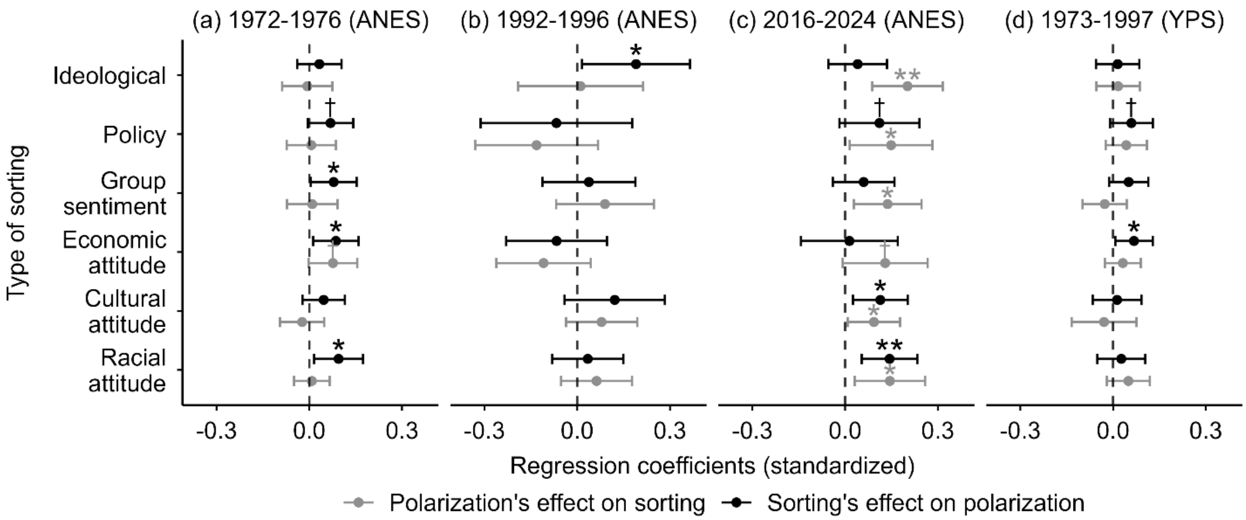
This figure tracks the items used to measure partisan sorting for all panels analyzed in Analysis #5a and Appendix G. Ideological sorting items are the same as those listed in Figure 2 for all panels except 1973–1997, which is limited to symbolic ideology and conservatives FT. FT=feeling thermometer.

Appendix M: A discussion of CLPMs and RI-CLPMs

Cross-lagged panel models (CLPMs) estimate causal relationships by assessing whether an initial observation of one variable predicts a later value of another variable, on the assumption that the prior variable caused the change (Finkel 1995). This is a common method of observational causal inference in political science (Carsey and Layman 2006; Highton and Kam 2011; Margolis 2018). However, they face increasing scrutiny for their inability to distinguish within-person effects from between-person effects: Changes in the dependent variable can come from measurement error or trait instability rather than the causal influence of the lagged variable. This may lead to spurious causal effects between correlated variables (Hamaker, Kuiper, and Grasman 2015; Lucas 2023).

Social scientists continue to debate the best way to improve on the CLPM. The most common option is the random-intercept cross-lagged panel model (RI-CLPM), which leverages at least three waves to account for variables' inter-wave stability, thereby separating between- and within-person effects (Lucas 2023). Unfortunately, RI-CLPMs suffer from separate issues, particularly when analyzing panels with only three waves measured at wide time intervals. First, RI-CLPMs require more power than standard CLPMs, increasing the risk of a Type II error. Second, as Bakker et al. (2021) demonstrate, RI-CLPMs can be twice as downwardly biased as traditional CLPMs are upwardly biased when the causal effect of a variable occurs far faster than the time between measurements (a likely issue here, given two- to nine-year gaps between waves).

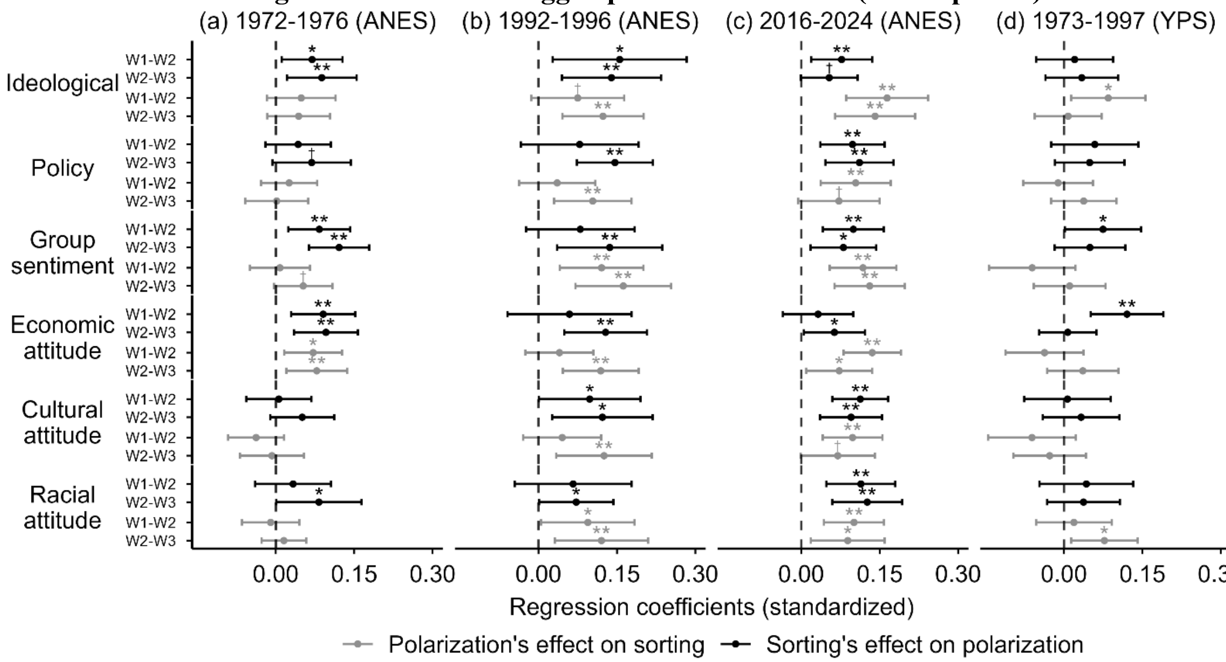
Figure OA.17: Reciprocal relationships between attitudinal sorting and affective polarization (random-intercept cross-lagged panel models)



Coefficients from random-intercept cross-lagged panel models applied to three ANES panels and the 1973–1997 waves of the Youth-Parent Socialization study (YPS). Appendix L lists items used to measure sorting. Models include demographic controls and, for 2016–2024, an indicator for wave 1's survey mode. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

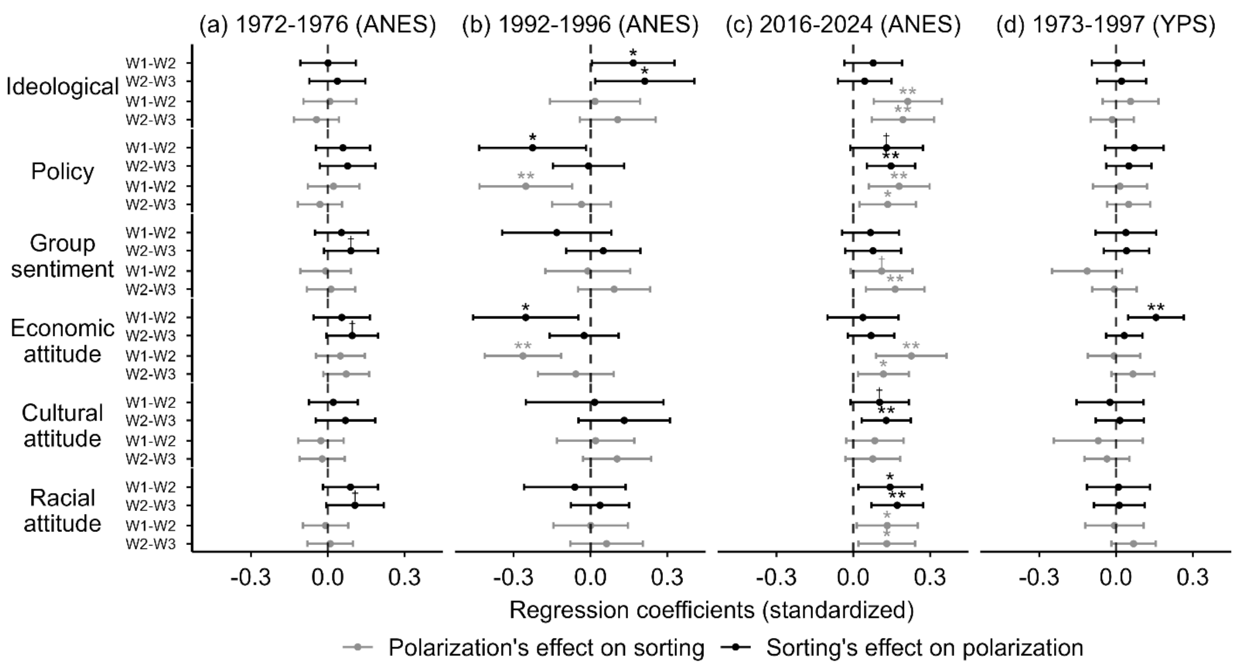
I apply RI-CLPMs to all four panels used in Analysis #5a (see Figure OA.17). The results are somewhat consistent with the CLPMs shown in Figure 7, particularly in the 1972–1976 and 2016–2024 panels. I find mainly null within-person effects for both sorting and polarization in the 1992–1996 panel, while the Youth-Parent Socialization Study RI-CLPMs indicate at least marginally significant influence for two forms of attitudinal sorting. Critically, only the significant effects from 2016–2024 consistently replicate across individual waves (see Figure OA.19).

Figure OA.18: Cross-lagged panel model results (wave-specific)



*Coefficients from cross-lagged panel models shown in Figure 7, but with results separated by wave. †p < .10, *p < .05, **p < .01.*

Figure OA.19: Random-intercept cross-lagged panel model results (wave-specific)



*Coefficients from random-intercept cross-lagged panel models shown in Figure OA.17, but with results separated by wave. †p < .10, *p < .05, **p < .01.*

However, Table OA.11 displays the results of power analyses for both causal pathways across all four panels. I find that only the 2016–2024 panel meets the conventional power threshold of 0.80 (and only for polarization, not partisan sorting). This suggests causal pathways could be overlooked by underpowered analyses of the other panels. Nevertheless, these RI-CLPM estimates—particularly outside of the 2016–2024 panel—should be interpreted cautiously.

Table OA.11: Power analyses for random-intercept cross-lagged panel models

Panel	Effect	N	ICC	β	SE	Bias	Coverage	Power
1972–1976	Sorting → Polarization	862	0.28	0.07	0.04	−0.00	0.95	0.36
	Polarization → Sorting	862	0.28	0.02	0.04	0.00	0.93	0.09
1992–1996	Sorting → Polarization	444	0.33	0.09	0.06	−0.00	0.94	0.26
	Polarization → Sorting	444	0.33	0.08	0.06	−0.00	0.95	0.30
2016–2024	Sorting → Polarization	1638	0.30	0.08	0.03	0.00	0.97	0.69
	Polarization → Sorting	1638	0.30	0.14	0.03	−0.00	0.94	0.98
1973–1997	Sorting → Polarization	592	0.12	0.04	0.04	−0.00	0.99	0.08
	Polarization → Sorting	592	0.12	0.03	0.05	−0.00	0.96	0.09

This table displays the results of Monte Carlo power analyses run on the RI-CLPMs displayed in Figure OA.17. Using the powRICLPM R package, I run 500 simulated datasets, with inputs—including intraclass correlation (ICC) and cross-lagged coefficients (β)—averaged across all six RI-CLPM fits for each panel.

Ultimately, a stricter test of causal relationships between attitudinal sorting and affective polarization simply requires better evidence, such as a panel that captures an exogenous shift in variables (e.g., Lenz 2012) or a panel with more waves and wide coverage of attitudes. In the meantime, this paper's panel analyses provide suggestive evidence, and the experiments in Analysis #5b corroborate the finding that attitudinal sorting has a causal effect on polarization.

Appendix N: Studies used in experimental meta-analysis

Table OA.12: Studies used in experimental meta-analysis

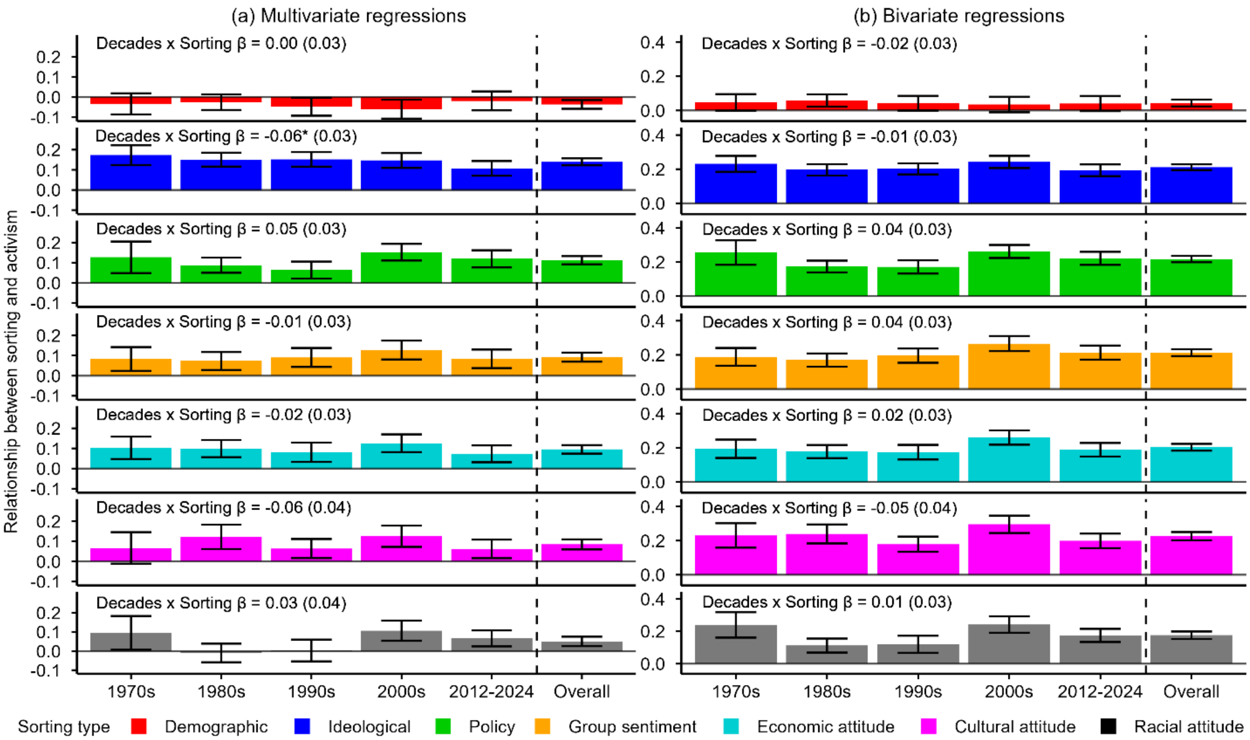
Study	Description	β
Druckman et al. (2022)	Subjects rated hypothetical out-partisans who randomly varied on both symbolic ideology and political engagement (e.g., "Conservative Republicans who frequently talk about politics"). I compare the difference in animus toward a) out-partisans with no ideological information and b) ideologically unsorted out-partisans.	−0.07 (0.01)
Dias and Lelkes (2022), Study 2	Subjects received three text vignettes describing hypothetical individuals. The vignettes randomly included information about party ID, policy preferences, or both. I compare the difference in animus toward out-partisans with a) no policy information vs. b) counter-stereotypical policy information.	−0.10 (0.02)
Lelkes (2021)	Subjects read about a hypothetical out-party candidate with a randomized set of moderate or extreme issue positions (or none). I compare the difference in animus toward candidates with a) no positions vs. b) moderate positions.	−0.05 (0.01)
Orr and Huber (2020), Study 1	Subjects received three text vignettes describing a hypothetical individual. The vignettes randomly included information about party ID, policy preferences, or both. I compare the difference in animus toward out-partisans with a) no policy information vs. b) counter-stereotypical policy information.	−0.04 (0.02)
Orr and Huber (2020), Study 2	See above, except subjects received nine text vignettes.	−0.05 (0.01)
Orr and Huber (2020), Study 3	See above, except subjects received one text vignette.	−0.06 (0.02)
Rogowski and Sutherland (2016)	Subjects evaluated two hypothetical candidates running for Congress who were randomly placed either close together or far apart on a (visual) ideological scale. The paper analyzes the effect of convergent vs. divergent candidates on differences in respondents' feeling thermometer ratings.	−0.11 (0.01)

This table describes the experiments used in Analysis #5b (see Figure 8). Rogowski and Sutherland's (2016) β and standard error are taken from Table 1, Column 1. All other estimates are produced from replication materials.

Appendix O: Beyond affective polarization

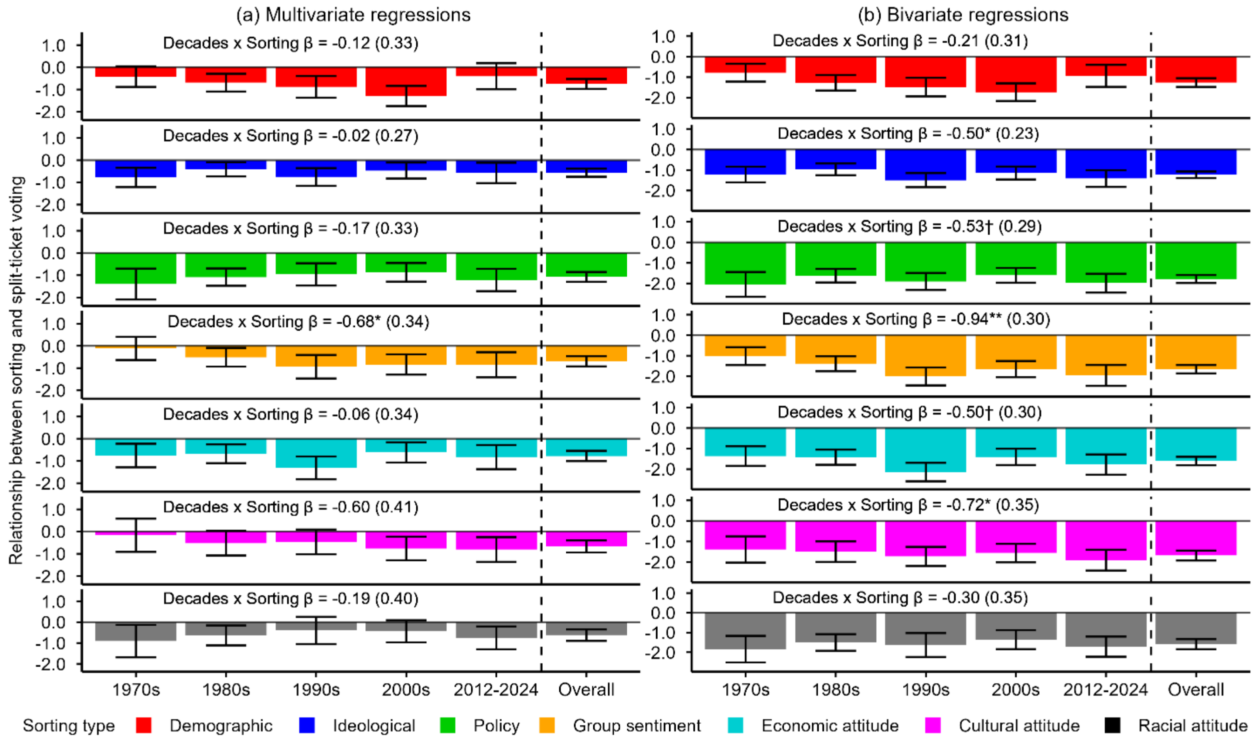
I replicate Analysis #2 with three measures of political behavior replacing feeling thermometer polarization as the dependent variable: an additive five-item index of political activism (Figure OA.20), self-reported split-ticket voting (Figure OA.21), and validated vote turnout (Figure OA.22). I also include two proxies for partisan social identity based on standard seven-point party ID: stability (Figure OA.23) and strength (Figure OA.24). Folded party ID strength is not a perfect proxy for partisan social identity, but it has been used by prior work (e.g., Mason 2015; Mason and Wronski 2018) and is the only option for over-time analyses.

Figure OA.20: Partisan sorting's over-time relationships with political activism



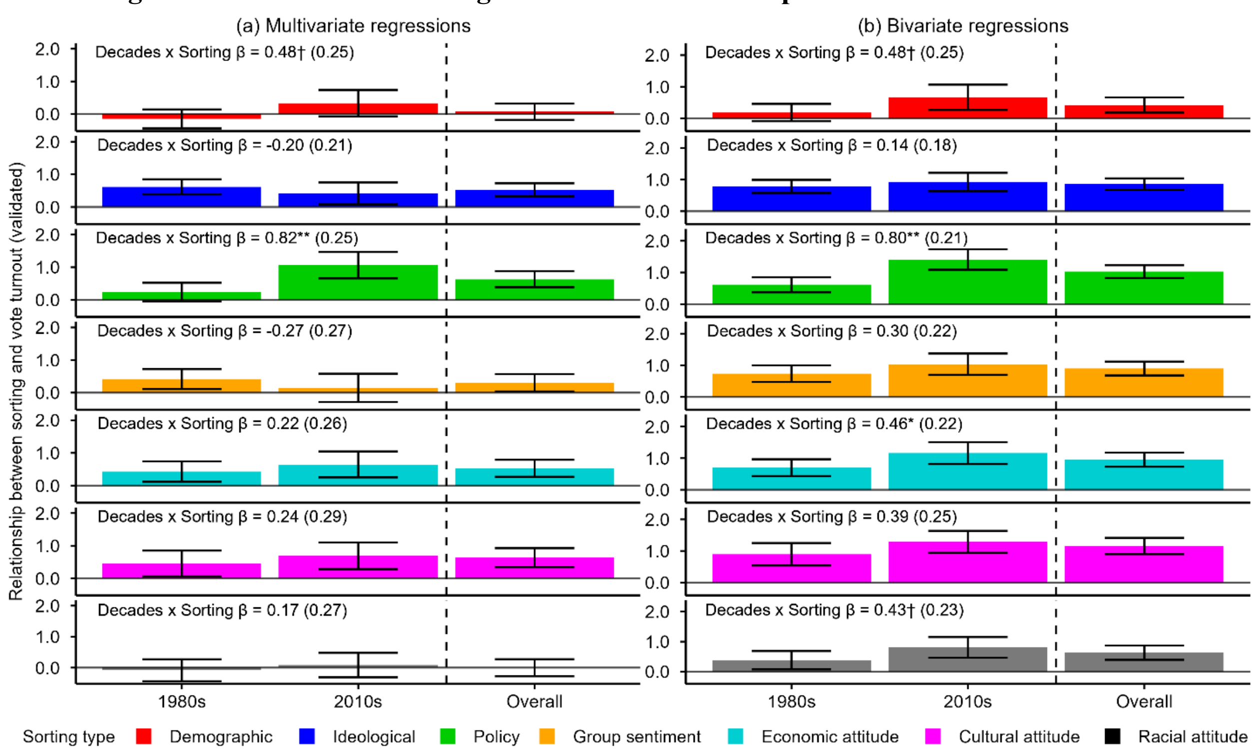
In this figure, I replicate Analysis #2 (Figure 5) with activism as the dependent variable. Activism is measured using a five-item additive scale of self-reported political activity: trying to influence others' vote, attending a political rally, working for a party or candidate, displaying a campaign button or sticker, and donating to a party or candidate. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

Figure OA.21: Partisan sorting's over-time relationships with split-ticket voting



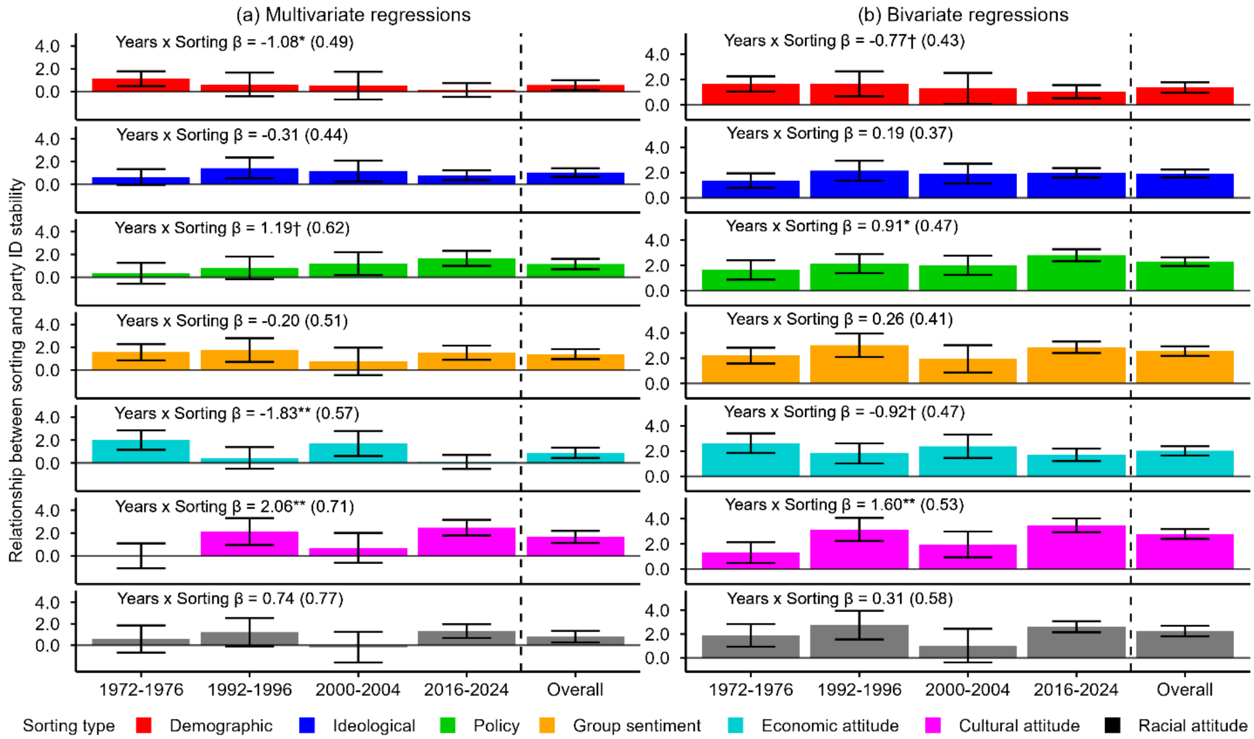
*In this figure, I replicate Analysis #2 (Figure 5) with split-ticket voting as the dependent variable in probit regressions (1=self-reported split ticket). †p < .10, *p < .05, **p < .01.*

Figure OA.22: Partisan sorting's over-time relationships with validated vote turnout



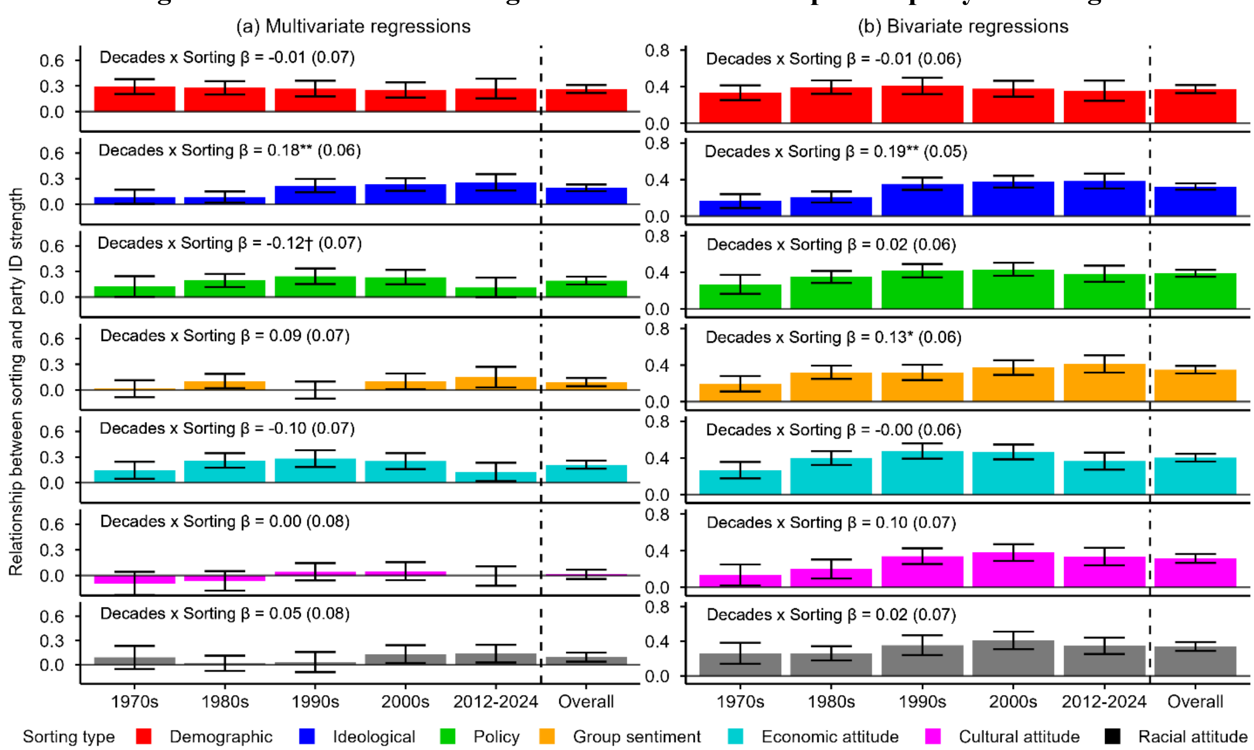
*In this figure, I replicate Analysis #2 (Figure 5) with validated vote turnout as the dependent variable in probit regressions (1=confirmed turnout). This analysis subsets to the years where the ANES verified respondent turnout using the voter file: 1980, 1984, 1988, 2012, and 2016. †p < .10, *p < .05, **p < .01.*

Figure OA.23: Partisan sorting's over-time relationships with party ID stability



*In this figure, I replicate Analysis #2 (Figure 5) with party ID stability as the dependent variable in probit regressions (1=stable three-point party ID, 0=party ID changes in final wave). This analysis subsets to ANES respondents in 1972, 1992, 2000, and 2016 who participated in the final waves of their respective panels. †p < .10, *p < .05, **p < .01.*

Figure OA.24: Partisan sorting's over-time relationships with party ID strength



*In this figure, I replicate Analysis #2 (Figure 5) with folded party ID strength as the dependent variable (1=strong partisans, 0.5=weak partisans, and 0=partisan leaners). †p < .10, *p < .05, **p < .01.*

References

Bailey, Michael A. 2024. *Polling at a Crossroads: Rethinking Modern Survey Research.* Cambridge: Cambridge University Press.

Bakker, Bert N., Yphtach Lelkes, and Ariel Malka. 2021. “Reconsidering the Link Between Self-Reported Personality Traits and Political Preferences.” *American Political Science Review* 115 (4): 1482–1498.

Bonica, Adam. 2018. “Inferring Roll-Call Scores from Campaign Contributions Using Supervised Machine Learning.” *American Journal of Political Science* 62 (4): 830–848.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. [1960] 1980. *The American Voter.* Chicago: University of Chicago Press.

Carsey, Thomas M., and Geoffrey C. Layman. 2006. “Changing Sides or Changing Minds? Party Identification and Policy Preferences in the American Electorate.” *American Journal of Political Science* 50 (2): 464–477.

Cavari, Amnon, and Guy Freedman. 2018. “Polarized Mass or Polarized Few? Assessing the Parallel Rise of Survey Nonresponse and Measures of Polarization.” *Journal of Politics* 80 (2): 719–725.

Cavari, Amnon, and Guy Freedman. 2023. “Survey Nonresponse and Mass Polarization: The Consequences of Declining Contact and Cooperation Rates.” *American Political Science Review* 117 (1): 332–339.

Clinton, Josh, Jennifer Agiesta, Megan Brennan, Camille Burge, Marjorie Connelly, Ariel Edwards-Levy, Bernard Fraga, et al. 2021. “2020 Pre-Election Polling: An Evaluation of the 2020 General Election Polls.” Technical Report, *American Association for Public Opinion Research*. https://aapor.org/wp-content/uploads/2022/11/Task-Force-2020-Primaries-Oct-6-2020-002.pdf.

Dias, Nicholas C., and Yphtach Lelkes. 2022. “The Nature of Affective Polarization: Disentangling Policy Disagreement from Partisan Identity.” *American Journal of Political Science* 66 (3): 775–790.

Druckman, James N., Samara Klar, Yanna Krupnikov, Matthew Levendusky, and John Barry Ryan. 2022. “(Mis)estimating Affective Polarization.” *Journal of Politics* 84 (2): 1106–1117.

Finkel, Steven E. 1995. *Causal Analysis with Panel Data.* Thousand Oaks, CA: Sage Publications.

Hamaker, Ellen L., Rebecca M. Kuiper, and Raoul P. P. P. Grasman. 2015. “A Critique of the Cross-Lagged Panel Model.” *Psychological Methods* 20 (1): 102–116.

Hartman, Erin, and Melody Huang. 2024. “Sensitivity Analysis for Survey Weights.” *Political Analysis* 32 (1): 1–16.

Highton, Benjamin, and Cindy D. Kam. 2011. "The Long-Term Dynamics of Partisanship and Issue Orientations." *Journal of Politics* 73 (1): 202–215.

Horan, Patrick M. 1971. "Social Positions and Political Cross-Pressures: A Re-Examination." *American Sociological Review* 36 (4): 650–660.

Kaufman, Aaron R., and Jon C. Rogowski. 2024. "Presidential Policymaking, 1877–2020." *Political Science Research and Methods* 12 (4): 687–705.

Lelkes, Yphtach. 2021. "Policy over Party: Comparing the Effects of Candidate Ideology and Party on Affective Polarization." *Political Science Research and Methods* 9 (1): 189–196.

Lenz, Gabriel S. 2012. *Follow the Leader? How Voters Respond to Politicians' Policies and Performance.* Chicago: University of Chicago Press.

Lucas, Richard E. 2023. "Why the Cross-Lagged Panel Model Is Almost Never the Right Choice." *Advances in Methods and Practices in Psychological Science* 6 (1): 1–22.

Margolis, Michele F. 2018. "How Politics Affects Religion: Partisanship, Socialization, and Religiosity in America." *Journal of Politics* 80 (1): 30–43.

Mason, Lilliana. 2015. "'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science* 59 (1): 128–145.

Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity.* Chicago: University of Chicago Press.

Mason, Lilliana, and Julie Wronski. 2018. "One Tribe to Bind Them All: How Our Social Group Attachments Strengthen Partisanship." *Political Psychology* 39 (S1): 257–277.

Mellon, Jonathan, and Christopher Prosser. 2021. "Correlation with Time Explains the Relationship between Survey Nonresponse and Mass Polarization." *Journal of Politics* 83 (1): 390–395.

Mellon, Jonathan, and Christopher Prosser. 2025. "Regularized Regression Can Reintroduce Backdoor Confounding: The Case of Mass Polarization." *American Political Science Review* 119 (4): 2002–2010.

Montgomery, David. 2019. "CityLab Congressional Density Index." GitHub. https://github.com/theatlantic/citylab-data/tree/master/citylab-congress (accessed September 4, 2025).

Muchlinski, David, David Siroky, Jingrui He, and Matthew Kocher. 2016. "Inferring Roll-Call Scores from Campaign Contributions Using Supervised Machine Learning." *Political Analysis* 24 (1): 87–103.

NYU Langone Health. 2022. "2022 CDHD District Density Index." Congressional District Health Dashboard. https://www.congressionaldistricthealthdashboard.org/2022-cdhd-district-density-index (accessed September 27, 2025).

Orr, Lilla V., and Gregory A. Huber. 2020. "The Policy Basis of Measured Partisan Animosity in the United States." *American Journal of Political Science* 64 (3): 569–586.

Peress, Michael. 2010. "Correcting for Survey Nonresponse Using Variable Response Propensity." *Journal of the American Statistical Association* 105 (492): 1418–1430.

Peyton, Kyle. 2020. "Does Trust in Government Increase Support for Redistribution? Evidence from Randomized Survey Experiments." *American Political Science Review* 114 (2): 596–602.

Rogowski, Jon C., and Joseph L. Sutherland. 2016. "How Ideology Fuels Affective Polarization." *Political Behavior* 38: 485–508.

Thornton, Judd R. 2025. "The Impact of Survey Mode on the Measurement of Polarization in the Public: 2016–2020 U.S. Panel Evidence." *American Politics Research* 53 (5): 531–543.

Tyler, Matthew, and Shanto Iyengar. 2024. "Testing the Robustness of the ANES Feeling Thermometer Indicators of Affective Polarization." *American Political Science Review* 118 (3): 1570–1576.

Tyler, Matthew, D. Sunshine Hillygus, Matthew DeBell, Ted Brader, Shanto Iyengar, Daron Shaw, and Nicholas A. Valentino. 2025. "Why are surveys struggling to estimate vote shares?" *American Journal of Political Science* 1–15. https://doi.org/10.1111/ajps.70051.

van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2011. "mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45 (3): 1–67.