There Is Only One Question: Diffuse Institutional Confidence in a Polarized America

Sean Kates* Jonathan M. Ladd† Joshua Tucker‡

January 7, 2026

Abstract

Public confidence in institutions shapes compliance, cooperation, and the acceptance of political and quasi-political outcomes. Yet confidence batteries are ordinal, multi-domain, and often politicized, raising the risk that apparent trends and group gaps are artifacts of measurement, or misleading about policy implications. In this paper, we assess confidence in public and private American institutions during the period from 2018 to 2023, treating "institutional confidence" as a potential multi-sectoral measure of legitimacy while evaluating change both across parties and time. Using a three-wave panel with partial re-interviews and a series of ordered-categorical bifactor confirmatory factor analyses, we identify a general confidence factor that spans both public and private institutions, alongside smaller domain-specific currents. Multi-group invariance tests support within-wave party comparisons on a common latent scale and reveal large, persistent partisan differences: Democrats score substantially higher on generalized institutional confidence, with gaps ranging from 0.38 to 0.79 standard deviations. Leveraging a common-item core and partial scalar invariance across waves, we estimate a moderate decline in generalized confidence of about 0.27 standard deviations over the study period. Finally, multiple-indicator multiple-cause (MIMIC) differential item functioning screens show that a limited set of institutions remains uniquely politicized net of generalized confidence and domain structure—including national media, the executive, political parties, and the FBI. Together, these findings imply that legitimacy in the contemporary United States is multi-sectoral and polarized, and that broad appeals to simply "restore trust" are unlikely to travel uniformly across institutional domains and populations. We conclude by offering a practical template for reporting institutional confidence on a genuinely comparable scale in polarized settings.

*University of Pennsylvania
†Georgetown University
‡New York University

Introduction

Public confidence in institutions is a cornerstone of democratic governance and policy capacity. When citizens believe that authorities and rule-makers are competent, fair, and acting in the public interest, they are more willing to comply with rules, accept adverse outcomes, and cooperate with collective projects (Easton 1975; Levi and Stoker 2000; Tyler 2006; Hetherington 2005; Devine 2024). When confidence erodes, policy capacity weakens and democratic accountability can become brittle: citizens attribute poor outcomes to corruption rather than error, treat losses as illegitimate rather than routine, and interpret procedural constraints as partisan weapons rather than neutral guardrails (Hetherington 1998; Norris 2011; Achen and Bartels 2016; Daniller 2016; Citrin and Stoker 2018). These stakes make "trust" and "confidence" popular topics in political science and adjacent fields.

Yet the object of citizen confidence is no longer well-described by "government" alone. Worldwide, citizens experience authority in a mixed system that includes elected officials and public agencies, but also major platform firms, banks, universities, media organizations, and other private or quasi-private institutions that exercise governance-like power (Cutler et al. 1999; Hall and Biersteker 2002; Büthe and Mattli 2011; Culpepper 2011; Gillespie 2018; Gorwa 2019). These actors write and enforce rules, limit and sanction behavior, and shape access to information and opportunity. Social media platforms moderate speech and coordinate "private ordering" of public discourse (Klonick 2018; Balkin 2018; Gillespie 2018; Gorwa 2019); credit rating agencies and financial institutions condition state capacity and distributive politics (Sinclair 2005; Fourcade and Healy 2017); universities structure life chances and social hierarchies (Collins 1979). If legitimacy is partly the public's belief that rule-makers are entitled to rule and will do so on acceptable terms, then institutional confidence is increasingly about legitimacy across *both* public and private authority, not only political institutions in the narrow sense.

Crisis of Trust?

Taking this broader view of legitimacy helps sharpen an empirical puzzle central in recent literature: are advanced democracies experiencing a "crisis" of trust, or merely institution- and period-specific fluctuations? Valgarðsson et al. (2025) argue that much of the apparent disagreement in the "crisis vs. trendless fluctuations" debate reflects analytical confusion: whether trust is declining depends on *which* institutions, *where*, and *when*. Their global synthesis (1958–2019) underscores heterogeneity across institutional targets and across national contexts, while also taking seriously the idea that aggregate legitimacy can falter in some places and periods (van der Meer 2017; Thomassen and van Ham 2017; Zmerli and van der Meer 2017). While they target a largely European sample, their general argument is especially suggestive for the United States in the late 2010s and early 2020s: a period marked by intense polarization, crisis politics, contested information environments, and salient conflicts over the perceived neutrality of both public institutions (e.g., executive agencies, courts, law enforcement) and private gatekeepers (e.g., social media and technology platforms, universities, media) (Hetherington and Rudolph 2015; Ladd 2012; Gillespie 2018; Balkin 2018).

Research Questions and Challenges

This raises a question that can be substantively important but methodologically delicate: when survey respondents report "confidence" in a long roster of institutions (both public and private) are they reporting a single underlying legitimacy judgment about institutional authority in general, or a patchwork of domain-specific evaluations? And if confidence is partly "general," are observed group gaps and time trends *real* differences in legitimacy, or artifacts of measurement? These issues arise because institutional confidence batteries are (i) ordinal, (ii) multi-domain, and (iii) politicized. The same response category can reflect different latent propensities across groups, and the same item can shift in meaning across time as party conflict reframes what the institution "stands for" (Millsap 2011; Meredith 1993; Vandenberg and Lance 2000). Without attention to comparability, scholars can mistakenly attribute differences in thresholds or loadings to substantive change, or miss genuine change when measurement drift cancels it out (Chen 2007; Rutkowski and Svetina 2014; Breustedt 2018; Schneider 2017). Each of these methodological concerns must be addressed while attempting to identify how institutional trust is structured in modern American politics.

Approach

Our paper responds to this conceptual and measurement challenge by treating institutional confidence as legitimacy across a mixed ecology of public and private authority, and by estimating it on a truly comparable latent scale. We analyze a three-wave U.S. panel with partial re-interviews spanning 2018–2023, modeling ordinal institutional confidence items using ordered-categorical confirmatory factor analysis (CFA) with a bifactor structure: a general factor G capturing what institutions share, plus smaller domain currents separating state/political institutions from private/civic authority. Bifactor modeling is well-suited to settings where a dominant general trait coexists with systematic domain clustering, allowing the analyst to avoid double-counting the general signal while still representing meaningful residual structure (Reise 2012; Rodriguez et al. 2016; Eid et al. 2017).

To make credible claims about partisan differences and time trends, we foreground measurement comparability rather than treating it as an afterthought. For within-wave party comparisons, we use multi-group invariance tests (configural/metric/scalar for ordered-categorical indicators) as the basis for interpreting latent mean gaps (Meredith 1993; Millsap 2011; Byrne et al. 1989; Chen 2007). This approach is now standard in the measurement literature and has been used in political-trust research to show that invariance cannot be assumed across countries or regime types (Breustedt 2018; Schneider 2017). For over-time inference, where exact invariance is often harder to sustain in politicized batteries, we adopt a conservative common-item strategy and implement partial scalar invariance on the intersection of institutions observed across waves (Vandenberg and Lance 2000; Rutkowski and Svetina 2014). Finally, to identify where politicization exceeds what the latent factors capture, we use multiple-indicator multiple-cause (MIMIC) differential item functioning screens—a well-established approach to detecting item-level group sensitivity conditional on latent traits (Jöreskog and Goldberger 1975; Woods 2009; Jessee 2021; Shiraito et al. 2023).

Summary of Findings

Substantively, our analyses yield three central findings. First, institutional confidence in this period is strongly structured by a general factor that spans both public and private institutions, consistent with the idea that many Americans answer these questions through a broad legitimacy lens rather than wholly institution-by-institution reasoning. Second, within waves, Democrats score substantially higher than Republicans on this general factor, with gaps on the order of 0.38 to 0.79 SDs, even after establishing comparability of the measurement model. Third, on a common-item core that supports over-time comparison, general confidence declines moderately—about 0.27 SDs from 2018 to 2023—suggesting that whatever separates partisans, there may also be a broader downward drift affecting both sides. The item-level DIF screens locate residual politicization where contemporary conflict is most salient: national media, the executive and parties, and institutions such as the FBI, with parallel sensitivity for prominent platform and knowledge institutions in some specifications.

Implications

Taken together, these results imply that restoring trust is not just a matter of "fixing politics" in the narrow sense. The trust problem is multi-sectoral: it combines a general legitimacy component, a large partisan divide, and a set of politicized institutions that function as symbolic battlegrounds. Methodologically, our contribution is a practical, replicable template for measuring and reporting institutional confidence on a comparable latent scale in polarized settings, combining ordinal bifactor measurement, explicit invariance logic, and transparent item-level diagnostics. Substantively, our contribution is to show that the legitimacy signal in confidence batteries is broader than state institutions alone, sharply stratified by party, and plausibly declining across the full electorate in the late 2010s and early 2020s.

The remainder of the paper proceeds as follows. We first describe the panel and the institutional battery, before presenting the ordered-categorical bifactor CFA framework, invariance strategy, and MIMIC DIF diagnostics. The Results section reports model fit, variance partition, partisan gaps, over-time change on the common-item core, and the item-level politicization map, before we conclude by discussing implications for legitimacy in mixed authority systems and for best practices in reporting institutional confidence.

Data and Methods

Design and Analytic Sample

In order to answer our questions, we analyzed the American Institutional Confidence Poll, a three-wave U.S. survey panel spanning two presidential cycles (2018-2023). Each nationally representative wave asked respondents a battery of questions about their confidence in a wide range of institutions, as well as their general support for democracy and various democratic norms.[1] Institutional-confidence items were measured on a 4-category ordered response scale, with respondents asked if they had "A great deal of confidence", "Only some confidence","Hardly any confidence", or "No confidence" in each institution. In our analyses, we limit analytic comparisons to respondents identifying with the two major parties (Democrats and Republicans) because the central inferential tasks are (i) establishing whether an institutional-confidence battery supports a comparable latent scale within a wave, and (ii) estimating party differences on that shared latent scale. Across waves, the Democrat/Republican sample sizes are: Wave 1 ($N_{\text{Dem}} = 2286$, $N_{\text{Rep}} = 1521$), Wave 2 ($N_{\text{Dem}} = 1810$, $N_{\text{Rep}} = 1152$), and Wave 3 ($N_{\text{Dem}} = 1701$, $N_{\text{Rep}} = 1187$).

Measures: Confidence in Public and Private Authority

Following our conceptualization of institutional confidence as legitimacy across *public and private authority*, we organize items into two substantive domains:

- **STATE/Political Authority** (10 items): the Executive, Congress, political parties, courts, state government, local government, FBI, military, national press/media, local police.

[1] The waves were collected in June/July 2018 (Wave 1), August/September 2021 (Wave 2), and October 2023 (Wave 3).

- **PRIVATE/Civic Authority** (9 items): Facebook, Google, Amazon, major companies, banks, colleges/universities, nonprofits, labor unions, religious organizations.

This partition is not intended to treat public and private authority as separate constructs; rather, it provides a theoretically interpretable way to model structured residual clustering after accounting for a general legitimacy-like evaluation.

Approach: Ordered-Categorical Bifactor CFA

Because confidence responses are ordered categories rather than interval data, we fit ordered-categorical CFA models in which each observed response $y_{ij} \in \{1, 2, 3, 4\}$ is generated by an underlying continuous propensity y^*_{ij} cut by thresholds τ:

$$y_{ij} = k \text{ if } \tau_{k-1} < y^*_{ij} \leq \tau_k, \qquad k = 1, \ldots, 4.$$

Models are estimated using Weighted Least Squares with Mean and variance Adjustment (WLSMV) with a probit link and θ parameterization, a standard approach for ordinal CFA with potentially skewed response distributions (Muthén 1984; Flora and Curran 2004). Our primary model is an ordered-categorical *bifactor* CFA (Reise 2012; Eid et al. 2017). All items load on a general factor G representing generalized institutional confidence (legitimacy across authority). Each item additionally loads on exactly one domain factor (STATE or PRIVATE). We impose orthogonality between G and each domain factor ($G \perp$ STATE and $G \perp$ PRIVATE), while allowing the two domain factors to correlate. For ease of interpretation, we place the latent factors on a standardized scale (mean and variance set by convention), so estimated differences can be read in standard deviation units.

Consistent with recommended practice (Brown 2015), we allow residual covariances for item pairs that share salience or framing beyond the factor structure. This helps avoid misattributing local dependence to the latent factors. In our analyses, we allow this for the three strongest pairs:

$$\text{Congress} \leftrightarrow \text{Parties}, \qquad \text{Google} \leftrightarrow \text{Amazon}, \qquad \text{Local police} \leftrightarrow \text{Military}.$$

Analysis I: Setting Up Competing Models

To assess whether the battery supports a dominant generalized confidence factor spanning public and private authority, we compare the bifactor model to a correlated two-factor alternative whcih loads items on STATE and PRIVATE only groupings only. Model fit is summarized with the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR). Beyond fit, we summarize how much common variance is attributable to the general factor versus the domains using explained common variance (ECV) from the bifactor solution (Rodriguez et al. 2016). ECV(G) captures the share of common variance explained by G; ECV(STATE) and ECV(PRIVATE) analogously capture domain shares.

Analysis II: Assessing Generalized Trust Over Time

A central challenge in studying trends in institutional confidence is that batteries often change across waves (items are added, dropped, or reworded). If one compares latent means using different sets of indicators, apparent "change" may partly reflect changes in measurement content rather than real shifts in attitudes. To avoid that confound, we estimate over-time change using a harmonized *common-item core*: the subset of confidence items asked with comparable wording and response options in all three waves.

Step 1: Fit the same measurement model across waves. On the common-item core, we estimate a multi-group (multi-wave) ordered-categorical bifactor CFA in which every item loads on the general factor G and on its designated domain factor (STATE or PRIVATE), with the same limited set of theory-motivated residual covariances included in each wave. We use WLSMV with a probit link and the θ parameterization to respect the ordered response scale (Muthén 1984; Flora and Curran 2004). This step imposes the same *form* of the measurement model across waves (the same items indicate the same latent factors), establishing a coherent baseline for comparability (Meredith 1993; Vandenberg and Lance 2000).

Step 2: Impose scalar comparability across waves. To interpret changes in latent means over time, we require that the measurement relationship between items and the latent construct is stable across waves. For ordered-categorical indicators, scalar invariance corresponds to equality constraints on factor loadings (so the latent metric is comparable) and on item thresholds (so response categories map to the underlying response propensities in the same way across waves) (Meredith 1993; Millsap 2011; Steenkamp and Baumgartner 1998). In our common-core model, these scalar constraints are supported across waves, implying that observed differences in responses can be attributed to differences in latent confidence rather than changes in how respondents use the response categories.

Step 3: Identify and interpret wave means on a common scale. We anchor the latent scale by fixing the Wave 1 mean of G to zero and scaling G so that Wave 1 provides the reference metric. This makes the estimated Wave 2 and Wave 3 means interpretable as deviations from Wave 1 in Wave 1 latent standard deviation units. We then estimate wave-to-wave contrasts (e.g., Wave 3 minus Wave 1) as linear combinations of model parameters, with standard errors obtained from the model-implied covariance matrix.

This procedure yields an estimate of over-time change in generalized institutional confidence that is (i) based on the same measurement content in every wave and (ii) identified under full scalar comparability across waves. In the Results, we report the latent mean of G by wave on the common core, along with the Wave 3–Wave 1 contrast and its uncertainty.

Analysis III: Within-Wave Party Comparability Using Measurement Invariance

Within each wave, we test whether Democrats and Republicans can be compared on a common latent confidence scale using multi-group measurement invariance (MI) tests (Meredith 1993; Millsap 2011; Vandenberg and Lance 2000).[2] The baseline measurement model is our ordered-categorical bifactor CFA: every item loads on the general factor G, each item addi-

[2]For our purposes, "invariance" means that the survey items relate to the latent construct in the same way across groups, so that group differences can be interpreted as differences in latent confidence rather than differences in measurement.

tionally loads on its assigned domain factor (STATE or PRIVATE), and the same small set of theory-motivated residual covariances is included in both groups.

We then fit three models that impose progressively stronger *across-party* equality constraints:

1. **Configural invariance (same measurement *form*):** We require that the *structure* of the model is the same for Democrats and Republicans. The same items indicate the same latent factors (all items load on G and on their designated domain), the same residual pairs are allowed, and each ordinal item is represented by the same number of cutpoints. However, the *values* of loadings and thresholds are freely estimated within each party. This tests whether the same conceptual measurement model is plausible in both groups.

2. **Metric-G invariance (equal *relationships* to G):** Building on the configural model, we constrain the factor loadings on the general factor G to be equal across Democrats and Republicans. Intuitively, this asks whether each item is equally diagnostic of generalized confidence in both parties (i.e., whether a one-unit change in G implies the same expected shift in each item's latent response propensity). We leave domain-factor loadings free, allowing the STATE/PRIVATE residual structure to differ by party.

3. **Scalar-G invariance (equal *thresholds* for the response categories):** For ordered-categorical items, scalar invariance is implemented by constraining the item thresholds to be equal across groups. Threshold equality means that choosing (for example) "A great deal of confidence" versus "Only some confidence" corresponds to the same location on the item's underlying response propensity for Democrats and Republicans. With equal cutpoints (and metric-G constraints), differences in observed responses can be attributed to differences in the latent means rather than differences in how groups use the response categories.

We evaluate invariance using practical changes in approximate fit, treating MI as supported when $\Delta\text{CFI} \geq -0.010$ and $\Delta\text{RMSEA} \leq 0.015$ when moving to stricter constraints (Cheung and Rensvold 2002; Chen 2007; Rutkowski and Svetina 2014).

Analysis IV: Identifying Residual Politicization Using MIMIC DIF screens

Finally, we map item-level politicization beyond generalized confidence using MIMIC DIF screens (Jöreskog and Goldberger 1975). Within each wave, we estimate direct effects of party on items while controlling for G and domain factors. Significant direct effects indicate items whose partisan differences exceed what is accounted for by generalized confidence and domain structure. We use MIMIC DIF as a diagnostic complement to the multi-group CFA above: while the latter establishes that party comparisons of the latent factor are valid, MIMIC DIF pinpoints which specific institutions exhibit additional partisan separation beyond generalized confidence and domain structure (Woods 2009; Jessee 2021).

Results

Analysis I: Model Fit

We begin by assessing whether institutional confidence behaves primarily as evaluations of two distinct sectors (STATE versus PRIVATE) or as a broader legitimacy-like judgment that spans both. Table 1 compares a correlated two-factor model to an ordered-categorical bifactor specification by wave. In all three waves, the bifactor model fits substantially better (CFI near 0.99; RMSEA near 0.06), indicating that responses are not well captured by two correlated domains alone. Instead, a single generalized confidence factor accounts for a large share of covariation across the entire roster of institutions, with domain factors capturing structured deviations.

Fit improvements alone do not determine whether G is substantively dominant. Table 2 reports explained common variance (ECV), which partitions common variance across G and domain factors. ECV(G) rises from 0.199 in Wave 1 to 0.470 in Wave 3, implying that nearly half of the shared variation in institutional ratings in the final wave is attributable to a single generalized confidence factor. Domain factors remain meaningful, but their relative share declines over time, consistent with a measurement environment in which respondents increasingly apply a general legitimacy-like lens across both public and private authority.

Table 1: Model fit by wave (ordered-categorical CFA, WLSMV/probit).

Wave	Model	CFI	RMSEA	SRMR
1	Two-factor (STATE, PRIVATE)	0.940	0.131	0.104
1	Bifactor (G + domains)	0.989	0.060	0.045
2	Two-factor (STATE, PRIVATE)	0.966	0.122	0.096
2	Bifactor (G + domains)	0.992	0.062	0.045
3	Two-factor (STATE, PRIVATE)	0.968	0.117	0.092
3	Bifactor (G + domains)	0.993	0.058	0.044

Notes: Approximate fit indices shown.

Table 2: Explained common variance (ECV) from the bifactor model.

Wave	ECV(G)	ECV(STATE)	ECV(PRIVATE)
1	0.199	0.450	0.350
2	0.411	0.328	0.261
3	0.470	0.296	0.234

Notes: ECV(G) is the share of common variance attributable to the general factor; ECV(STATE) and ECV(PRIVATE) are the corresponding shares attributable to the domain factors (Rodriguez et al. 2016).

Analysis II: Generalized Trust over Time

To assess over-time change, we fit the model on a harmonized common-item core and compare wave means on a comparable scale with Wave 1 anchored at 0. Table 3 shows a sharp decline from Wave 1 to Wave 2 and similarly low levels in Wave 3. Relative to Wave 1, the Wave 3 mean is -0.267 SD (SE = 0.029; 95% CI $[-0.323, -0.210]$), indicating a small-to-moderate downward shift in generalized confidence between 2018 and 2023, although this shift was already visible in 2021's Wave 2.

Table 3: Latent G means over time on the common-item core (Wave 1 reference).

Wave	Mean	SE	95% CI
1 (ref.)	0.000	–	–
2	-0.279	0.029	[-0.336, -0.222]
3	-0.267	0.029	[-0.323, -0.210]

Notes: Wave 1 is the reference group (latent mean fixed to 0). Estimates are in Wave 1 latent SD units.

Analysis III: Within-Wave Party Comparability

We next evaluate whether Democrats and Republicans can be compared on a common scale *within each wave*. Table 4 summarizes the multi-group invariance tests. Configural models fit well, and imposing our metric-G constraints produces only small changes in approximate fit. Scalar-G constraints also satisfy standard criteria in all waves, supporting the interpretation that partisan differences in latent G reflect differences on a comparable response scale rather than threshold drift or differential interpretation (Meredith 1993; Millsap 2011; Chen 2007).

Table 4: Within-wave measurement invariance by party (Democrat vs. Republican).

Wave	Config. CFI	Config. RMSEA	ΔCFI (Metric-G)	ΔRMSEA (Metric-G)	ΔCFI (Scalar-G)	ΔRMSEA (Scalar-G)
1	0.990	0.057	-0.004	0.007	0.000	-0.003
2	0.994	0.052	-0.002	0.007	0.000	-0.002
3	0.992	0.056	-0.003	0.006	0.000	-0.003

Notes: Metric-G constrains only G loadings equal across party; domain loadings are free. Scalar-G additionally constrains ordered thresholds equal across party. Traditional guidelines for plausible invariance are $\Delta\text{CFI} \geq -0.010$ and $\Delta\text{RMSEA} \leq 0.015$ (Cheung and Rensvold 2002; Chen 2007; Rutkowski and Svetina 2014).

Given scalar comparability, Table 5 reports latent differences in G, with Republicans as the within-wave reference group. Democrats score higher in every wave, and the gap grows across the period: 0.382 SD (Wave 1), 0.668 SD (Wave 2), and 0.786 SD (Wave 3). Because G spans the full roster of our institutions, these differences are best interpreted as broad-based gaps in generalized legitimacy rather than disagreement about any one institution.

Table 5: Democrat–Republican difference in latent G within wave (scalar MG-CFA).

Wave	Diff (Dem–Rep)	SE	95% CI
1	0.382	0.034	[0.315, 0.448]
2	0.668	0.040	[0.590, 0.746]
3	0.786	0.040	[0.707, 0.865]

Notes: Differences are within-wave latent SD units under `std.lv=TRUE`. Interpretability relies on within-wave scalar invariance (Table 4).

Because thresholds vary across items, latent SD units do not translate mechanically into a fixed number of response categories. The magnitudes here are therefore best interpreted as *broad shifts across many items simultaneously.* In that sense, partisan gaps of 0.38–0.79 SD reflect substantial differences in generalized confidence across the institutional roster. The over-time decline of approximately -0.27 SD is smaller than the partisan divide but still meaningful: it represents a detectable downward drift in generalized legitimacy spanning both public and private authority.

Analysis IV: Residual Politicization

Finally, we use MIMIC DIF screens to identify institutions whose partisan differences persist beyond the generalized confidence gap and domain structure. Table 6 reports the largest residual effects by wave (full set in Appendix Table 7). Residual partisan sensitivity concentrates in a consistent cluster: national political institutions (executive, Congress, parties), the national media, universities, major platforms, and the FBI. This implies that polarization in institutional confidence has both a broad component (generalized legitimacy) and a concentrated component in high-salience institutions that remain distinctive sites of partisan conflict.

Implications and Conclusions

Our substantive findings carry three implications for how scholars and practitioners should interpret institutional confidence in polarized settings. First, institutional confidence behaves like legitimacy across multiple sectors of authority. Across waves, a single generalized factor increasingly explains shared variation in confidence judgments. This supports treating institutional confidence batteries as more than a summary of attitudes toward government: they can reflect perceived legitimacy of authority that spans public institutions and major private governors. In practice, this means that debates about "trust in institutions" are plausibly debates about authority and legitimacy in a mixed governance environment, where private platforms, universities, firms, and banks can function as consequential rule-setters alongside state actors.

Second, polarization is both about levels of trust, and specific institutional battlegrounds. Even after accounting for generalized confidence and domain structure, residual partisan

Table 6: Top residual partisan sensitivity (MIMIC DIF), by wave.

Wave	Item	$\hat{\beta}$	SE
1	National media/press	0.631	0.036
	Colleges/universities	0.524	0.036
	Political parties	0.497	0.036
	Labor unions	0.481	0.036
	FBI	0.461	0.036
2	Executive branch	0.813	0.041
	National media/press	0.793	0.041
	Congress	0.752	0.041
	Political parties	0.744	0.042
	Colleges/universities	0.697	0.041
3	Executive branch	0.929	0.042
	National media/press	0.885	0.042
	Colleges/universities	0.824	0.042
	FBI	0.791	0.042
	Political parties	0.793	0.043

Notes: $\hat{\beta}$ is the direct effect of party on the item conditional on G and the domain factors (items $\sim$ party | factors). Positive values indicate higher confidence among Democrats than Republicans net of latent confidence. Full item-level results are in Appendix Table 7.

sensitivity repeatedly concentrates in a consistent cluster (executive/parties/Congress, national media, universities, major platforms, FBI). This suggests that some institutions have become especially salient sites for partisan contestation: evaluations of these institutions diverge beyond what would be expected from a general "trustful vs. distrustful" disposition alone.

Third, the combination of a broad decline and a large partisan divide complicates practical strategies to "restore trust." The over-time decline indicates a generalized downward drift, while the widening party gap indicates asymmetric legitimacy across coalitions. Where institutional confidence supports compliance, cooperation, and the acceptance of authoritative decisions, such patterns imply practical strain on governance capacity and dispute resolution (Tyler 2006; Levi and Stoker 2000; Hetherington 2005; Devine 2024). For policymakers and institutional leaders, the key lesson is that interventions limited to one sector (e.g., only "government trust") may miss a broader legitimacy environment; and interventions that ignore institution-specific politicization may fail even when generalized confidence is addressed.

From a methodological standpoint, the results underscore a reporting template suited to polarized environments: (i) model confidence as ordered-categorical, (ii) establish within-wave comparability before interpreting group gaps, (iii) separate generalized confidence from domain structure, and (iv) disclose where institution-specific politicization remains via DIF screens. This approach preserves substantively interpretable comparisons while making the measurement assumptions explicit.

Note for SPSA Readers:

Three scope conditions are worth emphasizing. First, the analysis relies on a three-wave panel with partial re-interviews; panel retention and item nonresponse can affect representativeness and may interact with trust itself. Where population inference is central, weighted robustness checks and attrition diagnostics should accompany the main estimates. We are working on this, so hopefully this goes away for a future draft. Second, our over-time estimate is based on a harmonized common-item core; this improves comparability but may understate changes occurring in institutions that enter or exit the battery across waves. This is also only a minor concern for us, as we added very few institutions to each wave, but it is worth noting and perhaps discussing during the panel. Third, while the bifactor model provides a disciplined way to separate generalized confidence from domain structure and localized dependence, any latent-variable specification remains an approximation; alternative specifications (e.g., approximate invariance or Bayesian regularization) are valuable robustness tools we haven't completed yet.

Appendix

Table 7: Full MIMIC DIF results by wave (items $\sim$ party | factors).

Wave	Item	$\hat{\beta}$	SE	z
1	National media/press	0.631	0.036	17.547
	Colleges/universities	0.524	0.036	14.363
	Political parties	0.497	0.036	13.704
	Labor unions	0.481	0.036	13.337
	FBI	0.461	0.036	12.773
	Google	0.408	0.036	11.230
	Facebook	0.294	0.036	8.165
	Amazon	0.284	0.037	7.754
	Congress	0.267	0.036	7.456
	Executive branch	0.236	0.036	6.521
	Courts	0.231	0.036	6.466
	State government	0.228	0.036	6.329
	Local government	0.212	0.036	5.899
	Major companies	0.184	0.036	5.147
	Banks	0.156	0.036	4.375
	Military	0.138	0.036	3.848
	Local police	0.097	0.036	2.707
	Nonprofits	0.072	0.036	1.999
	Religious organizations	0.061	0.036	1.687
2	Executive branch	0.813	0.041	19.659
	National media/press	0.793	0.041	19.200
	Congress	0.752	0.041	18.213
	Political parties	0.744	0.042	17.863

Continued on next page

Wave	Item	$\hat{\beta}$	SE	z
	Colleges/universities	0.697	0.041	16.846
	Labor unions	0.661	0.041	15.976
	FBI	0.606	0.041	14.746
	Google	0.604	0.041	14.808
	Facebook	0.426	0.041	10.377
	Amazon	0.392	0.041	9.466
	Major companies	0.363	0.041	8.839
	Courts	0.326	0.041	7.967
	Banks	0.302	0.041	7.367
	State government	0.279	0.041	6.785
	Local government	0.251	0.041	6.112
	Nonprofits	0.176	0.041	4.299
	Military	0.164	0.041	4.001
	Local police	0.142	0.041	3.470
	Religious organizations	0.102	0.041	2.483
3	Executive branch	0.929	0.042	22.060
	National media/press	0.885	0.042	21.121
	Colleges/universities	0.824	0.042	19.396
	FBI	0.791	0.042	18.731
	Political parties	0.793	0.043	18.642
	Labor unions	0.717	0.042	17.114
	Google	0.645	0.042	15.305
	Congress	0.566	0.042	13.511
	Amazon	0.511	0.042	12.094
	Facebook	0.482	0.042	11.448
	Major companies	0.433	0.042	10.197

Continued on next page

Wave	Item	$\hat{\beta}$	SE	z
	Courts	0.402	0.042	9.521
	Banks	0.371	0.042	8.778
	State government	0.327	0.042	7.775
	Local government	0.296	0.042	7.040
	Military	0.226	0.042	5.371
	Local police	0.212	0.042	5.041
	Nonprofits	0.171	0.042	4.076
	Religious organizations	0.135	0.042	3.225

References

Achen, C. H. and Bartels, L. M. (2016). It feels like we're thinking: The rationalizing voter and electoral democracy. *PS: Political Science & Politics*, 49(1):4–8.

Balkin, J. M. (2018). Free speech is a triangle. *Columbia Law Review*, 118(7):2011–2055.

Breustedt, W. (2018). Testing the measurement invariance of political trust across the globe: A multiple group confirmatory factor analysis. *methods, data, analyses*, 12(1):7–46.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. Guilford Press, New York, 2 edition.

Büthe, T. and Mattli, W. (2011). *The New Global Rulers: The Privatization of Regulation in the World Economy*. Princeton University Press, Princeton, NJ.

Byrne, B. M., Shavelson, R. J., and Muthen, B. (1989). Structural equation modeling with lisrel, prelis, and prelis-pc. *Journal of Econometrics*, 41(1):181–211.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3):464–504.

Cheung, G. W. and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2):233–255.

Citrin, J. and Stoker, L. (2018). Political trust in a cynical age. *Annual Review of Political Science*, 21:49–70.

Collins, R. (1979). *The Credential Society: An Historical Sociology of Education and Stratification*. Academic Press, New York.

Culpepper, P. D. (2011). *Quiet Politics and Business Power: Corporate Control in Europe and Japan*. Cambridge University Press, Cambridge.

Cutler, A. C., Haufler, V., and Porter, T., editors (1999). *Private Authority and International Affairs*. State University of New York Press, Albany, NY.

Daniller, A. M. (2016). Can citizens care too much? investment in election outcomes and perceptions of electoral legitimacy. *Electoral Studies*, 44:151–161.

Devine, D. (2024). Does political trust matter? a meta-analysis on the consequences of trust. *Political Behavior*, 46:2241–2262.

Easton, D. (1975). A re-assessment of the concept of political support. *British Journal of Political Science*, 5(4):435–457.

Eid, M., Geiser, C., and Koch, T. (2017). A review of bifactor modeling: Theory and applications. *Psychological Methods*, 22(2):255–277.

Flora, D. B. and Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4):466–491.

Fourcade, M. and Healy, K. (2017). Seeing like a market. *American Journal of Sociology*, 122(6):1661–1705.

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press, New Haven, CT.

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6):854–871.

Hall, R. B. and Biersteker, T. J., editors (2002). *The Emergence of Private Authority in Global Governance*. Cambridge University Press, Cambridge.

Hetherington, M. J. (1998). The political relevance of political trust. *American Political Science Review*, 92(4):791–808.

Hetherington, M. J. (2005). *Why Trust Matters: Declining Political Trust and the Demise of American Liberalism*. Princeton University Press, Princeton, NJ.

Hetherington, M. J. and Rudolph, T. J. (2015). *Why Washington Won't Work: Polarization, Political Trust, and the Governing Crisis*. University of Chicago Press, Chicago.

Jessee, S. A. (2021). Differential item functioning in political science: A framework for assessing measurement invariance. *Political Analysis*, 29(3):380–396.

Jöreskog, K. G. and Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 70(351):631–639.

Klonick, K. (2018). The new governors: The people, rules, and processes governing online speech. *Harvard Law Review*, 131(6):1598–1670.

Ladd, J. M. (2012). *Why Americans Hate the Media and How It Matters*. Princeton University Press, Princeton, NJ.

Levi, M. and Stoker, L. (2000). Political trust and trustworthiness. *Annual Review of Political Science*, 3:475–507.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4):525–543.

Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. Routledge, New York.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49(1):115–132.

Norris, P. (2011). *Democratic Deficit: Critical Citizens Revisited*. Cambridge University Press, Cambridge.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5):667–696.

Rodriguez, A., Reise, S. P., and Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3):223–237.

Rutkowski, L. and Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, 74(1):31–57.

Schneider, I. (2017). Can we trust measures of political trust? assessing measurement equivalence in diverse regime types. *Social Indicators Research*, 133(3):963–984.

Shiraito, Y., Hsu, Y.-L., and Tingley, D. (2023). Survey satisficing and differential item functioning in cross-national research. *Political Analysis*, 31(2):223–243.

Sinclair, T. J. (2005). *The New Masters of Capital: American Bond Rating Agencies and the Politics of Creditworthiness*. Cornell University Press, Ithaca, NY.

Steenkamp, J.-B. E. M. and Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25(1):78–90.

Thomassen, J. and van Ham, C. (2017). Political trust and the crisis of democracy. In Zmerli, S. and van der Meer, T. W. G., editors, *The Handbook on Political Trust*, pages 476–493. Edward Elgar Publishing, Cheltenham.

Tyler, T. R. (2006). *Why People Obey the Law*. Princeton University Press, Princeton, NJ, 2 edition.

Valgarðsson, V., Jennings, W., Stoker, G., Bunting, H., Devine, D., McKay, L., and Klassen, A. (2025). A crisis of political trust? global trends in institutional trust from 1958 to 2019. *British Journal of Political Science*, 55:e15.

van der Meer, T. W. G. (2017). Political trust and the "crisis of democracy". Oxford Research Encyclopedia of Politics.

Vandenberg, R. J. and Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1):4–70.

Woods, C. M. (2009). Evaluation of mimic-model methods for dif testing with comparison to two-group analysis. *Applied Psychological Measurement*, 33(5):335–359.

Zmerli, S. and van der Meer, T. W. G., editors (2017). *The Handbook on Political Trust*. Edward Elgar Publishing, Cheltenham.