

Modeling Trends in Aggregate Public Opinion Using Multiple Survey Sources: An Overview of Approaches, Assumptions, and Tradeoffs

Piotr Koc, GESIS–Leibniz Institute for the Social Sciences, Germany

Marta Kołczyńska, Institute of Political Studies of the Polish Academy of Sciences, Poland

Submitted to *PS: Political Science and Politics* as part of the Symposium

Data Harmonization in Political Science and Related Fields.

Version prepared for the 2026 APSA conference in Boston, MA.

ABSTRACT

Due to the accumulation of data from cross-national survey projects as well as methodological advancements, data from surveys can be combined to estimate latent trends in mass public opinion. This study analyzes the proposed approaches to modeling trends in public opinion, identifies main conceptual and statistical issues related to applying these models, and proposes ways to address them in substantive research.

The past two decades have seen substantial developments in cross-national survey research. Existing survey programs have expanded to countries previously not included and new multi-wave cross-national survey projects have been launched. The resulting accumulation of survey data, combined with advances in statistical computing, makes it increasingly feasible to estimate trends in mass values and attitudes across time and space using multiple survey sources.

These developments open new research avenues because they enable analyses of the macro-level consequences of mass public opinion and they allow scholars to address fundamental questions in political science, such as whether support for democracy is necessary for democratic consolidation and whether trust in state institutions improves institutional performance. Research using latent trend models also may examine convergence and divergence in attitudes across societal groups, as well as features of opinion distributions beyond central tendencies, including ideological polarization.

This article reviews recent developments in modeling trajectories of mass public opinion using cross-national survey data and their substantive applications. It identifies key conceptual and statistical challenges and discusses strategies to address them, providing guidelines for rigorous future research in this area.

MODELING TRENDS IN PUBLIC OPINION

The challenge in estimating public opinion trajectories based on cross-national survey data lies primarily in the differences in measurement across survey projects and the limited availability of repeated items across countries and years. The model-based approaches to addressing these challenges proposed in the literature share basic principles but also differ in important ways.

One commonality is the focus on reflective constructs, as opposed to formative constructs (Singh, Heller, and Quandt 2026), and the reliance on item response theory (IRT) for estimating the latent means per country and year of attitudes measured with ordinal survey items. Another commonality is the workflow, which includes identifying survey items corresponding to the construct of interest, selecting surveys based on item availability (Hu, Tai, and Solt 2026), and relying on the same software (see the online appendix).

An important distinction among the proposed approaches is the level of analysis. Individual-level models estimate the latent trait for individual respondents; these estimates then can be aggregated to the group level to obtain the country–year–level trajectory (Kołczyńska et al. 2024). The most suitable measurement model is the cumulative ordinal model, which assumes that survey responses reflect the categorization of a continuous latent attitude into an observed ordinal variable. This model can be specified to address the different lengths of the response scales to the same question (Gangl 2023; Kołczyńska et al. 2024), which addresses one of the main challenges in harmonization.

Group-level models assume that the unobservable trait is a property of the group, not the individual (Mislevy 1983). In this approach, individual-level data are aggregated to the group level, by calculating counts or proportions, before being input to the model. Two modeling approaches are represented in the literature. The first maintains the ordinal character of the data and simply counts the number of observations for each response category to apply to an ordinal model (Caughey, O’Grady, and Warshaw 2019; Solt 2020). The second approach dichotomizes the attitudinal responses before aggregation and applies models for binary data (Claassen 2019).

Group-level models are faster to run¹ and tolerate more sparsity in the country–year–item matrix of coverage by input survey data, at the cost of making additional assumptions. One assumption is the independence of responses to different items from the same respondents—for example, as in the case of items on endorsement of different features of democracy included in the same survey. Ignoring the clustering of responses within respondents results in the underestimation of the uncertainty around group means (Caughey

¹ These models are computationally intensive and require days or weeks to run on a high-performance computing cluster. Computational feasibility is one of the challenges faced by researchers in this area.

and Warshaw 2015; Kołczyńska and Bürkner 2024). However, simulations suggest that credible interval coverage deteriorates only when the proportion of items per respondent is substantial (Berwick and Caughey 2025). Individual-level models reflect the data-generation process more faithfully and enable the direct verification of model assumptions. However, they require more computing power as well as more input data—each item must be repeated multiple times in each country, which is not required by group-level models.

These models are applied to survey items that were not necessarily designed to measure the same construct or had their measurement properties validated. A prominent application of the models includes investigation of the links between mass support for democracy and a country's quality of democratic governance (Claassen 2020a, 2020b).² This research used items that measure diffuse support for democracy from 14 cross-national survey projects administered to nationally representative samples of citizens from 150 countries between 1988 and 2017 (Claassen 2020a). The selection of items was based on their prior use in research on democratic support. In total, 10 types of items were selected, including support for forms of government other than democracy and assessment of the suitability of democracy for one's country. Each project contributed from one to seven items. This analysis relied on the group-level modeling approach with dichotomized items. The following sections use this research as an illustration.

CONCEPTUAL ISSUES

As previously mentioned, mass public-opinion models provide great opportunities. Yet, using these models comes with conceptual challenges. First, there is an inherent tension between maximizing country–year coverage and ensuring measurement validity. When constructing

² References to other applications are in the online appendix.

measures—whether to capture individual- or aggregated-level characteristics—the content of the construct must be considered carefully (Wilson 2023). Crucially, we must establish the conditions for something to be defined as “support for democracy” and, based on these conditions, identify distinct theoretical attributes that will be linked to the items. In developing these attributes, the issue of differentiation must be addressed—that is, how distinct the attributes of the construct are from those of other closely related constructs (Gerring 1999). If the attributes overlap substantially, it becomes difficult—if not impossible—to distinguish the resulting measures. Wholly operationalizing the substantive content of the construct is related directly to content validity (Jackman 2009). Attributes with linked items that enable differentiation between the construct (and its measure) and other similar constructs (and their measures) are key to establishing divergent validity. That is, it is a situation in which the correlation between measures of distinct constructs is low enough to support the claim that the measures are assessing different constructs (Rönkkö and Cho 2022).

The current practice apparently favors maximizing country–year coverage at the expense of measurement validity. Validation frequently is reduced to nomological validation, in which measures are assessed by correlating them with variables representing other constructs. The goal is to verify whether the observed sign and magnitude align with well-established expectations derived from theories involving the construct of interest and related constructs (Adcock and Collier 2001). Although these correlations may provide preliminary evidence of validity, they cannot substitute the previous conceptual steps. In practice, two similar yet distinct constructs often yield similar correlation patterns, making it impossible to determine for which construct the correlations provide validity evidence.

Second, this coverage-maximizing approach is based on the premise that constructs are universally comparable across countries and over time. This assumption may stem from the fact that some models (Claassen 2019; Solt 2020) allow for country-specific deviations in

item parameters (an issue discussed in the next section), thereby mitigating some of the negative consequences of measurement non-invariance. However, this does not address the broader question of whether the constructs can be compared meaningfully across diverse contexts without resorting to conceptual stretching (Sartori 1970; see also De Sio, Katsanidou, and Borghetto 2026). Although this question is not new, it becomes particularly pressing when models are used to make comparisons on an unprecedented scale, often spanning more than 100 countries and 30-plus years.

Sartori (1970) argued that rather than stretching concepts whose meanings are tied to a specific set of attributes and cases, we may need to “climb the ladder of abstraction” (or generality; see Collier and Mahon 1993). This involves selecting more general concepts, with a reduced set of properties, to enable analysis in more countries and time periods (Sartori 1970). Similar insights recently echoed in the psychological literature, which emphasizes the need to establish theoretical comparability before statistical tests because the latter cannot ensure the former (Fischer et al. 2023; Fischer and Rudnev 2024). It is important to note that climbing the ladder of abstraction to ensure theoretical comparability should not be dismissed as mere semantic refinement. It is necessary because the theoretical processes and mechanisms that underly one construct might differ fundamentally from those that underly another more general construct (Fischer and Rudnev 2024). For instance, the causes of support for political systems more broadly may not fully overlap with or carry the same weight as those that underlie support for democracy.

STATISTICAL ISSUES

The second category of issues concerns the assumptions that underly the latent trend models and the challenges of making inferences based on the resulting estimates. All of the models discussed in this analysis originate from the IRT family and rely on several assumptions, of

which the dimensionality and conditional (or local) independence assumptions are fundamental (Embretson and Reise 2000). The dimensionality assumption implies that item responses are driven by one or more latent variables, whereas conditional independence stipulates that—conditional on the latent variable(s)—item responses are statistically independent (De Ayala 2022). In applied research, dimensionality sometimes is reduced to a unidimensionality assumption—namely, that covariance among items is driven by a single latent variable (Ip 2010). This assumption should be understood as a modeling choice, typically motivated by considerations of parsimony, data availability, and computational resources. In practice, unidimensionality and local independence often are violated, but such violations do not necessarily render useless models that are affected by them. Rather, their implications for inference depend on the severity of the violations (De Ayala 2022), and unidimensional models with approximate local independence still can provide useful and interpretable summaries of latent trends.

In the Bayesian framework, violations of the conditional-independence assumption—which may be indicative of issues with unidimensionality (Yen 1993)—can be investigated by comparing observed data with posterior predictive data for item pairs, using odds ratios or correlation residuals as discrepancy measures (Fox 2010). This, however, is possible only when working with individual-level data because the joint distribution of item responses is needed to calculate the discrepancy measures.

A different approach—feasible also with group-level IRT models—is to use model comparison to determine whether multidimensionality is more plausible than assumed unidimensionality. To implement this approach, a multidimensional factor structure (Berwick and Caughey 2025)³ can be specified along with a unidimensional model. We then can

³ However, this requires justifying multiple factors and conducting validation checks.

compare the model fit using Pareto-smoothed importance sampling leave-one-out cross-validation and select the better-fitting model (Fujimoto and Falk 2024). Although comparing model fit cannot fully replace discrepancy measures for investigating local independence violations, which can arise for various reasons (Fox 2010, 125), it is preferable to no verification at all.

The second issue pertains to measurement invariance. As mentioned previously, some models attempt to address invariance by allowing for country-specific deviations in item parameters: intercepts⁴. However, allowing deviations only in intercepts assumes constant slopes across countries, which implies that the items have consistent meanings and equally strong relationships to the latent factor (Kline 2016). This would mean that, for example, (observed) support for military rule would be correlated equally strongly with (latent) support for democracy across countries that have and have not experienced one or the other form of governance, which is a strong assumption. Therefore, researchers should consider accounting for country-specific deviations in slopes (De Jong, Steenkamp, and Fox 2007). Not considering variations in all item parameters can result in underestimating the uncertainty around country–year trends (Koc and Pokropek 2022). Nevertheless, these adjustments presuppose fundamental comparability of the latent construct across contexts.

The third issue concerns incorporating uncertainty into explanatory models in which estimates of latent trends are used as dependent or independent variables. To date, four main approaches have been proposed. The first approach simply uses the posterior mean, ignoring uncertainty; however, this can lead to bias and underestimating uncertainty in regression coefficients (Stoetzer, Zhou, and Steenbergen 2024). The second approach—the composition method—was popularized in political science by Treier and Jackman (2008) and recently

⁴ For a thorough discussion of this approach, see Warncke and Carlin (2026).

applied by, among others, Tai, Hu, and Solt (2024) and Kołczyńska and Bürkner (2026). The method is simple to implement but relies on strong assumptions about the ignorability of covariates in estimating latent trends (Treier and Jackman 2008). The third approach, discussed by Claassen (2022), is a single-step estimation method similar to Structural Equation Modeling. The key advantage is that it relaxes the ignorability assumption made by the composition method; however, it might be difficult to implement because more parameters must be estimated in one attempt. The fourth approach is the error-in-variable model (Kołczyńska and Bürkner 2021), which provides greater flexibility than the composition method by enabling explicit modeling of the measurement error (e.g., via covariates). A promising yet underutilized technique is overimputation (Blackwell, Honaker, and King 2017). Until recently, applying overimputation to customized autoregressive models was difficult, but recent developments using Integrated Nested Laplace Approximations have enabled its application to more complex models (Skarstein, Martino, and Muff 2023). Despite this variety of methods, research comparing their advantages and disadvantages remains limited (but see Claassen 2022).

The fourth issue concerns making inferences. The purpose of the models discussed herein is to produce estimates at the country–year-level based on individual-level measures. These country–year estimates then are related to other group-level characteristics, even if only for the purpose of validation. If the relationship of interest exists at the individual rather than the group level, using aggregated group-level data may result in problematic interpretations that rely on strong ecological assumptions (Berwick and Caughey 2025). For example, if higher-income countries observe higher average levels of support for democracy than lower-income countries, it does not necessarily follow that higher-income individuals universally support democracy more than lower-income individuals—either within countries or overall in pooled individual-level data. Similarly, a micro-level association between political interest and

education does not necessarily translate into the same association between country averages. Therefore, making inferences using country–year estimates is justified when (1) the country-level aggregation can be defended on theoretical grounds—that is, the construct is conceptually meaningful as a country-level property; and (2) the relationships of interest are at the group level. Alternatively, when the primary interest lies in reconstructing individual-level relationships, additional ecological assumptions or statistical adjustments (Jiang et al. 2020) should be explicitly incorporated.

CONCLUSION

Models for estimating latent trends in values and attitudes represent an important methodological advancement, which opens new research avenues including longitudinal analyses of the macro-level consequences of public opinion that address foundational questions in political science. Like any method, however, these models must be applied with careful attention to their assumptions and to the suitability of the available data. This article highlights opportunities for improving studies of mass public-opinion trends by encouraging researchers to reflect on the comparability of concepts across countries and over time, to verify dimensionality and measurement invariance, and to account for measurement uncertainty in explanatory models.

Research on mass public-opinion trends critically depends on the availability of high-quality survey data, particularly from cross-national and multi-wave projects that provide similarly worded items across countries and over time. The quality of survey samples likewise is crucial for the reliability of resulting estimates; however, some deviations can be mitigated using techniques such as multilevel regression and post-stratification (Gelman and Little 1997; Park, Gelman, and Bafumi 2004), as applied by Caughey and Warshaw (2015) and Kołczyńska et al. (2024).

The models discussed in this article are relatively recent, and their statistical properties—including the consequences of violations of key assumptions—remain incompletely understood. Future research, particularly simulation studies, is needed to assess the sensitivity of estimates to departures from measurement invariance and to different patterns of sparsity in item–country–year coverage, such as items asked only in selected countries. Such research is essential for developing more concrete and robust guidelines for rigorous research on latent trends in mass public opinion.

CONFLICTS OF INTEREST

The authors declare that there are no ethical issues or conflicts of interest in this research.

ACKNOWLEDGMENTS

We thank the editors, guest co-editors, reviewers, and other symposium contributors for their feedback. Marta Kołczyńska acknowledges funding from the National Science Centre (Grant Nos. 2022/04/Y/HS6/00024 and 2019/32/C/HS6/00421).

REFERENCES

- Adcock, Robert, and David Collier. 2001. “Measurement Validity: A Shared Standard for Qualitative and Quantitative Research.” *American Political Science Review* 95 (3): 529–46. DOI:10.1017/S0003055401003100.
- Berwick, Elissa, and Devin Caughey. 2025. “MODGIRT: Multidimensional Dynamic Scaling of Aggregate Survey Data.” *Political Analysis* 1–16. DOI:10.1017/pan.2024.25.
- Blackwell, Matthew, James Honaker, and Gary King. 2017. “A Unified Approach to Measurement Error and Missing Data: Overview and Applications.” *Sociological Methods & Research* 46 (3): 303–41. DOI:10.1177/0049124115585360.
- Caughey, Devin, Tom O’Grady, and Christopher Warshaw. 2019. “Policy Ideology in European Mass Publics, 1981–2016.” *American Political Science Review* 113 (3): 674–93. DOI:10.1017/S0003055419000157.
- Caughey, Devin, and Christopher Warshaw. 2015. “Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model.” *Political Analysis* 23 (2): 197–211. DOI:10.1093/pan/mpu021.

- Claassen, Christopher. 2019. "Estimating Smooth Country–Year Panels of Public Opinion." *Political Analysis* 27 (1): 1–20. DOI:10.1017/pan.2018.32.
- Claassen, Christopher. 2020a. "Does Public Support Help Democracy Survive?" *American Journal of Political Science* 64 (1): 118–34. DOI:10.1111/ajps.12452.
- Claassen, Christopher. 2020b. "In the Mood for Democracy? Democratic Support as Thermostatic Opinion." *American Political Science Review* 114 (1): 36–53. DOI:10.1017/S0003055419000558.
- Claassen, Christopher. 2022. "Including Measurement Uncertainty in Time-Series, Cross-Sectional Analyses: The Case of Mood and Democracy." www.ssrn.com/abstract=3924934.
- Collier, David, and James E. Mahon. 1993. "Conceptual 'Stretching' Revisited: Adapting Categories in Comparative Analysis." *American Political Science Review* 87 (4): 845–55. DOI:10.2307/2938818.
- De Ayala, Rafael J. 2022. *The Theory and Practice of Item Response Theory*. Second edition. New York and London: Guilford Press.
- De Jong, Martijn G., Jan-Benedict E. M. Steenkamp, and Jean-Paul Fox. 2007. "Relaxing Measurement Invariance in Cross-National Consumer Research Using a Hierarchical IRT Model." *Journal of Consumer Research* 34 (2): 260–78. DOI:10.1086/518532.
- De Sio, Lorenzo, Alexia Katsanidou, and Enrico Borghetto. Forthcoming. "Concepts Before Code: A Theoretical and Procedural Framework for Data Harmonization." *PS: Political Science & Politics*.
- Embretson, Susan E., and Steven Paul Reise. 2000. *Item Response Theory for Psychologists*. New York: Psychology Press. DOI:10.4324/9781410605269.
- Fischer, Ronald, Johannes Alfons Karl, Markus Luczak-Roesch, and Larissa Hartle. 2023. *Why We Need to Rethink Measurement Invariance: The Role of Measurement Invariance for Psychological Science*. DOI:10.31234/osf.io/3f9ca.
- Fischer, Ronald, and Maksim Rudnev. 2024. "From MIsgivings to MIse-en-scène: The Role of Invariance in Personality Science." *European Journal of Personality* 39 (4): 662–73. DOI:10.1177/08902070241283081.
- Fox, Jean-Paul. 2010. *Bayesian Item Response Modeling: Theory and Applications*. New York: Springer.
- Fujimoto, Ken A., and Carl F. Falk. 2024. "The Accuracy of Bayesian Model Fit Indices in Selecting Among Multidimensional Item Response Theory Models." *Educational and Psychological Measurement* 84 (2): 217–44. DOI:10.1177/00131644231165520.
- Gangl, Markus. 2023. "A Generalized Ordered Logit Model to Accommodate Multiple Rating Scales." *Sociological Methods & Research* 54 (1): 315–50. DOI:10.1177/00491241231186655.
- Gelman, Andrew, and Thomas C. Little. 1997. "Poststratification into Many Categories Using Hierarchical Logistic Regression." *Statistical Methodology* 23 (2): 127–35.
- Gerring, John. 1999. "What Makes a Concept Good? A Criterial Framework for Understanding Concept Formation in the Social Sciences." *Polity* 31 (3): 357–93. DOI:10.2307/3235246.

- Hu, Yue, Yuehong Cassandra Tai, and Frederick Solt. Forthcoming. “A TAO for Data Wrangling: A Practical Routine for Getting Past the ‘Janitor Work.’” *PS: Political Science & Politics*.
- Ip, Edward Haksing. 2010. “Empirically Indistinguishable Multidimensional IRT and Locally Dependent Unidimensional Item Response Models.” *British Journal of Mathematical and Statistical Psychology* 63 (2): 395–416. DOI:10.1348/000711009X466835.
- Jackman, Simon. 2009. “Measurement.” In *The Oxford Handbook of Political Methodology*, ed. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, 119–51. Oxford, UK: Oxford University Press. DOI:10.1093/oxfordhb/9780199286546.003.0006.
- Jiang, Wenxin, Gary King, Allen Schmalz, and Martin A. Tanner. 2020. “Ecological Regression with Partial Identification.” *Political Analysis* 28 (1): 65–86. DOI:10.1017/pan.2019.19.
- Kline, Rex B. 2016. *Principles and Practice of Structural Equation Modeling*. New York: Guilford Press.
- Koc, Piotr, and Artur Pokropek. 2022. “Accounting for Cross-Country/Cross-Time Variations in Measurement Invariance Testing: A Case of Political Participation.” *Survey Research Methods* 16 (1): 79–96. DOI:10.18148/SRM/2022.V16I1.7909.
- Kołczyńska, Marta, and Paul-Christian Bürkner. 2021. *Political Trust as a Cause and Consequence of Democracy: Longitudinal Analysis of European Data*. DOI:10.31235/osf.io/79dth.
- Kołczyńska, Marta, and Paul-Christian Bürkner. 2024. “Modeling Public Opinion Over Time: A Simulation Study of Latent Trend Models.” *Journal of Survey Statistics and Methodology* 12 (1): 130–54. DOI:10.1093/jssam/smad024.
- Kołczyńska, Marta, and Paul-Christian Bürkner. 2026. “Does Political Trust Strengthen Democracy? A Cross-National Longitudinal Analysis.” *International Political Science Review*, January 31. DOI:10.1177/01925121251405191.
- Kołczyńska, Marta, Paul-Christian Bürkner, Lauren Kennedy, and Aki Vehtari. 2024. “Modeling Public Opinion Over Time and Space: Trust in State Institutions in Europe, 1989–2019.” *Survey Research Methods* 18 (1): 1–19. DOI:10.18148/SRM/2024.V18I1.8119.
- Mislevy, Robert J. 1983. “Item Response Models for Grouped Data.” *Journal of Educational Statistics* 8 (4): 271–88. DOI:10.3102/10769986008004271.
- Park, David K., Andrew Gelman, and Joseph Bafumi. 2004. “Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls.” *Political Analysis* 12 (4): 375–85. DOI:10.1093/pan/mp024.
- Rönkkö, Mikko, and Eunseong Cho. 2022. “An Updated Guideline for Assessing Discriminant Validity.” *Organizational Research Methods* 25 (1): 6–14. DOI:10.1177/1094428120968614.
- Sartori, Giovanni. 1970. “Concept Misformation in Comparative Politics.” *American Political Science Review* 64 (4): 1033–53. DOI:10.2307/1958356.
- Singh, Ranjit Konrad, Ayline Heller, and Markus Quandt. 2026. “Assessing and Improving Comparability of Formative Individual-Level Measurement Instruments.” *PS: Political Science & Politics*. This issue.

- Skarstein, Emma, Sara Martino, and Stefanie Muff. 2023. “A Joint Bayesian Framework for Missing Data and Measurement Error Using Integrated Nested Laplace Approximations.” *Biometrical Journal* 65 (8): 2300078. DOI:10.1002/bimj.202300078.
- Solt, Frederick. 2020. *Modeling Dynamic Comparative Public Opinion*. DOI:10.31235/osf.io/d5n9p.
- Stoetzer, Lukas F., Xiang Zhou, and Marco Steenbergen. 2024. “Causal Inference with Latent Outcomes.” *American Journal of Political Science* 69 (2): 624–40. DOI:10.1111/ajps.12871.
- Tai, Yuehong Cassandra, Yue Hu, and Frederick Solt. 2024. “Democracy, Public Support, and Measurement Uncertainty.” *American Political Science Review* 118 (1): 512–18. DOI:10.1017/S0003055422000429.
- Treier, Shawn, and Simon Jackman. 2008. “Democracy as a Latent Variable.” *American Journal of Political Science* 52 (1): 201–17. DOI:10.1111/j.1540-5907.2007.00308.x.
- Warncke, Philip, and Ryan E. Carlin. 2026. “Bad Mood Rising? Assessing Scalar Invariance Violations with Comparative Democratic Support Data.” *Public Opinion Quarterly*. DOI:10.1093/poq/nfag025.
- Wilson, Mark R. 2023. *Constructing Measures: An Item Response Modeling Approach*. Second edition. New York and London: Routledge.
- Yen, Wendy M. 1993. “Scaling Performance Assessments: Strategies for Managing Local Item Dependence.” *Journal of Educational Measurement* 30 (3): 187–213. DOI:10.1111/j.1745-3984.1993.tb00423.x.

Appendix

Applications

In this section, we list several political science studies, in chronological order, that have analyzed latent trends in public opinion. Although not exhaustive, the list provides an overview of the themes addressed and the opportunities created by this research strategy.

Anthony J. McGann. 2014. “Estimating the Political Center from Aggregate Data: An Item Response Theory Alternative to the Stimson Dyad Ratios Algorithm.” *Political Analysis* 22(1): 115-129.

Devin Caughey and Christopher Warshaw. 2018. “Policy Preferences and Policy Change: Dynamic Responsiveness in the American States, 1936–2014.” *American Political Science Review* 112(2): 249-266.

Devin Caughey, Tom O’Grady, and Christopher Warshaw. 2019. “Policy Ideology in European Mass Publics, 1981–2016.” *American Political Science Review* 113(3): 674-693.

Parrish Bergquist and Christopher Warshaw. 2019. “Does Global Warming Increase Public Concern about Climate Change?” *The Journal of Politics* 8 (2): 686-691.

Christopher Claassen. 2020a. “Does Public Support Help Democracy Survive?” *American Journal of Political Science* 64(1): 118-134.

Christopher Claassen. 2020b. “In the Mood for Democracy? Democratic Support as Thermostatic Opinion.” *American Political Science Review* 114(1): 36-53.

Christopher Claassen and Lauren McLaren. 2022. “Does Immigration Produce a Public Backlash or Public Acceptance? Time-Series, Cross-Sectional Evidence from Thirty European Democracies.” *British Journal of Political Science* 52(3): 1013-1031.

Byung-Deuk Woo, Lindsey A. Goldberg, and Frederick Solt. 2023. “Public Gender Egalitarianism: A Dataset of Dynamic Comparative Public Opinion toward Egalitarian Gender Roles in the Public Sphere.” *British Journal of Political Science* 53(2): 766-775.

Marta Kołczyńska, Paul-Christian Bürkner, Lauren Kennedy, and Aki Vehtari. 2024. “Modeling Public Opinion Over Time and Space: Trust in State Institutions in Europe, 1989-2019.” *Survey Research Methods* 18(1): 1-19.

Chengyuan Ji, Junyan Jiang, and Yujin Zhang. 2024. “Political trust and government performance in the time of COVID-19.” *World Development* 176:106499.

Yue Hu and Frederick Solt. 2025. “Macrointerest Across Countries.” *British Journal of Political Science* 55:e71.

Marta Kołczyńska and Paul-Christian Bürkner. 2025. *Does political trust strengthen democracy?* Accessed September 2, 2025. https://osf.io/wd5hp_v2.

Viktor Valgarðsson et al. 2025. “A Crisis of Political Trust? Global Trends in Institutional Trust from 1958 to 2019.” *British Journal of Political Science* 55:e15.

Marta Kołczyńska and Paul-Christian Bürkner. 2026. “Does political trust strengthen democracy? A cross-national longitudinal analysis.” *International Political Science Review*. DOI: 10.1177/01925121251405191.

Software

The political science applications of latent trend models to estimating trajectories of public opinion rely on Stan, a probabilistic programming language for Bayesian data analysis, typically via Rstan or cmdstanr interfaces to R. Some studies use the R package brms, which enables users to use Stan using a user-friendly formula syntax rather than Stan code.

Paul-Christian Bürkner. 2017. “brms: An R Package for Bayesian Multilevel Models Using Stan.” *Journal of Statistical Software* 80(1): 1-28.

Paul-Christian Bürkner. 2018. “Advanced Bayesian Multilevel Modeling with the R Package brms.” *The R Journal* 10(1): 395-411.

Paul-Christian Bürkner. 2021. “Bayesian Item Response Modeling in R with brms and Stan.” *Journal of Statistical Software* 100(5): 1-54

Jonah Gabry et al. 2025. *cmdstanr: R Interface to 'CmdStan'*. R package version 0.9.0, <https://discourse.mcstan.org>.

R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Stan Development Team. 2025. *RStan: the R interface to Stan*. R package version 2.32.7. <https://mc-stan.org/>.